2012

13C-Metabolic flux analysis of soybean somatic embryos for identification of metabolic control points in developing seed

Quyen Xuan Truong
Iowa State University

13C-Metabolic flux analysis of soybean somatic embryos for identification of metabolic control points in developing seed

by

Quyen Xuan Truong

A dissertation submitted to the graduate faculty

in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Major: Chemical Engineering

Program of Study Committee:
Jacqueline V. Shanks, Major Professor
Julie A. Dickerson
Basil Nikolau
Laura R. Jarboe
Ian C. Schneider

Iowa State University
Ames, Iowa
2013

TABLE OF CONTENTS

Page

ABSTRACT iii

CHAPTER 1 INTRODUCTION AND LITERATURE REVIEW **1**

Metabolic Flux Analysis 2
13C-MFA in Plant Systems 9
Literature Review: soybean 11
Organization of Thesis 17
References 18

CHAPTER 2 INFLUENCE OF CARBON-TO-NITROGEN RATIOS ON SOYBEAN SOMATIC EMBRYOS (CV. JACK) GROWTH AND COMPOSITION **26**

Abstract 26
Introduction 27
Materials and Methods 30
Results 38
Discussion 45
Acknowledgements 52
References 52
Figure captions and Tables 55
Supplemental Data 61

CHAPTER 3 METABOLIC FLUX MAP COMPARISONS OF WILD-TYPE SOYBEAN (CV. JACK) ZYGOTIC AND SOMATIC EMBRYOS **70**

Abstract 70
Introduction 71
Materials and Methods 74
Results 87
Discussion 100
Acknowledgements 109
References 109
Figure captions and Tables 120
Supplement I: Figure and Tables 129
Supplement II: Detailed Calculation 146
Supplement III: Comparison of pathway models 150
Supplement IV: Isotopomer Balance of Rubisco reaction 159
Supplement V: Modify network with Transcripts and measurement data 161
Supplement VI: Parallel comparison of soybean zygotic studies 163
Supplement VII: Detailed procedure of labeling samples analysis 170

CHAPTER 4 SOYBEAN SOMATIC EMBRYOS METABOLIC FLUX MAP COMPARISON OF PLASTIC PHOSPHOGLUCOMUTASE KNOCK OUT AND DIFFERENT CARBON NUTRITIONAL SUPPLEMENT IN THE MEDIUM.....**205**

Abstract 205
Introduction 206
Materials and Methods 209
Results 221
Discussion 236
Acknowledgements 241
References 242
Figure captions and Tables 253
Supplement Data: Figure and Tables.....268

CHAPTER 5 CONCLUSIONS.....**305**

Conclusions.....305

ABSTRACT

Of the world's major agronomic food crops soybeans rank highest in protein content (~40% Dwt) while also containing significant quantities of oil (~20% Dwt). Based on these unique characteristics soy has become a mainstay in world agriculture, providing a protein source for livestock and human nutrition (68% of global vegetable protein meal consumption in 2011) as well as a primary source of vegetable oil (28% of global vegetable oil consumption in 2011; http://www.soystats.com/2012). The value of soy is therefore found both in its oil and protein content and increasing the content of both is therefore desirable. Research aimed at increasing the oil content, while leaving the protein content unchanged, has exposed a fundamental lack of understanding of resource partitioning and the factors that influence protein and oil content in the soybean seed. Somatic embryos have proven to be a highly productive platform for testing gene combinations designed to change soybean composition and provide a useful model for performing physiological and biochemical studies.

Soybean somatic embryos cultured in Soybean Histodifferentiation and Maturation (SHaM) medium were examined for their suitability as a model system for developing an understanding of assimilate partitioning and metabolic control points for protein and oil biosynthesis in soybean seed. It was postulated that at media compositions that were sufficient to support maximal growth rates, changes in the C:N ratio are likely to influence the partitioning of resources between the various storage products, especially protein and oil. As postulated, at steady-state growth rates embryo protein content was strongly correlated to decreasing C: N ratios and

increasing glutamine consumption rates. However, oil content remained relatively unchanged across the C: N ratio range tested and resources were instead directed towards the starch and residual biomass (estimated by mass balance) pools in response to increasing C: N ratios. Protein and oil were inversely related only at media sucrose concentrations below 88 mM, where carbon limited growth and no starch was found to accumulate in the tissues.

This work describes the in-depth studies of zygotic and SHaM embryos under similar culture conditions and carbon and nitrogen sources. There is no significantly different in relative growth rate for both embryos. Both protein and oil content were lower for SHaM embryos than in zygotic embryos; however, starch contents were comparable, and the balance of the biomass differences, which was accounted for by the residual (structural carbohydrate) pool, was higher in the SHaM embryos. Flux analysis in cultured embryos resulted changes in nitrogen uptake and flux into oil biosynthesis, respiratory flux (CO_2), glutamine biosynthesis flux, fluxes in the total of plastic and cytosolic of triose phosphate to phosphoenolpyruvate pathway, as well as an increase in tricarboxylic acid cycle activity for zygotic embryos. However, fluxes into structure and non-structure carbohydrates were significantly higher in SHaM embryos. Despite these differences, the NMR relative intensities of proteinogenic amino acids and labeling patterns of protein and starch-related glucosyl units were comparable between the two embryo types. Carbon labeling patterns of SHaM embryos well fitted with the metabolic network model of zygotic embryos with three compartments: cytosol, plastid, and mitochondrion. The observations described here

shed light onto metabolic pathways of SHaM embryos, especially as compared to soybean seed.

This thesis describes experiments in which we have used Metabolic Flux Analysis to investigate the influence of transgenic perturbations and nutritional status on resource partitioning in a soybean somatic embryo system. SHaM embryos of transgenic cultures with the plastidic phosphoglucomutase (PGM) gene knocked out (PGM-KO), and the control (PGM-null) are cultured in sucrose concentrations ranged from 88 to 234 mM as a carbon source and initial glutamine concentrations ranged from 20 to 60 mM as a nitrogen source. These concentrations correspond to C:N ratios ranging from 8.8 to 70.2. Two C: N mole ratio conditions are further examined through metabolic flux analysis with labeling experiment of U-$^{13}C_{12}$ sucrose for both PGM culture. The result indicates that: (1) protein and oil of PGM-KO were consistently higher than the PGM-null; (2) content in PGM-KO shows nearly two fold as compared to PGM-null; and (3) for both PGM culture, protein content is strongly correlated with the glutamine uptake rate. Fluxes through cytosolic glucose-6-phosphate isomerase, transketolase, and transaldolase, contributed significantly to the soluble sugar content for PGM-KO culture. These fluxes changed in response to the absence of starch synthesis.

CHAPTER 1

INTRODUCTION AND LITERATURE REVIEW

Soybean [*Glycine max* (L.) Merr.] seeds are valued for their protein and oil content, and when processed, provide the largest global source of vegetable protein for animal feed. Soybean seeds have a very high protein content (40%), but uniquely they are also high in oil (20%) (Sinclair and de Wit, 1975). The global consumption of vegetable oil has approximately doubled from the year 2000 to 2012, whereas the consumption of protein meal only increased by 150% (http://www.fas.usda.gov/psdonline/circulars/oilseeds.pdf). Increasing global demands for vegetable oil (for human nutrition; as a renewable source of energy, biodiesel; and as a feedstock for green chemistry), along with the increasing abundance of alternative cheap vegetable protein sources for livestock feed, may soon result in an era when the total value (oil production x price of oil/mass) of soybean oil becomes greater than that of the soybean protein. The conundrum is that the oil and protein content of soybean seeds are inversely related (Hartwig and Kilen, 1991), specifically a 1% reduction in oil content will lead to a 2% increase in protein content (Clemente and Cahoon, 2009). Although carbon shifts between protein and oil content, the underlying metabolic control mechanisms are not fully understood, and strong metabolic links between oil and storage protein synthesis are not apparent (Schwender *et al.*, 2003). The overall challenge is to increase the oil content of soybeans at the expense of seed components such as starch, soluble sugars, and cell wall material, while maintaining protein content at a level that the market demands;

i.e., a dehulled, defatted meal with 47.5 - 49% protein (http://www.soymeal.org/composition). Biotechnology and metabolic engineering seem promising to accomplish this challenging task (Lu *et al.*, 2011); however, in order to make progress we need to fully understand the control points of carbon partitioning which dictate the inverse relationship between protein and oil contents in developing soybean seed.

With the development of modern ‘‘-omics’’ techniques such as genomics, transcriptomics, proteomics, and metabolomics, along with methodologies of systems biology gaining popularity among both engineers and biologists, the wealth of information available on an increasing number of organisms is advantageous for metabolic engineering (Durot *et al.*, 2009; Feist *et al.*, 2009; Medini *et al.*, 2008). Metabolic engineering has been defined as ‘‘the directed improvement of product formation or cellular properties through the modification of specific biochemical reactions or introduction of new ones with the use of recombinant DNA technology’’(Stephanopoulos, 1999). Metabolomics aim to quantitatively measure concentrations of every metabolite within a cell population at a certain time, (Kell, 2004; Nielsen and Oliver, 2005; van der Werf *et al.*, 2007) while metabolic flux analysis (MFA) aims to quantify the flow of primarily carbon throughout a metabolic network. As a result, metabolic fluxes are the functional output of the transcriptome, proteome, and metabolome.

Metabolic Flux Analysis

Metabolic flux analysis (MFA) is one of the major tools available in metabolic engineering to generate quantitative metabolic flux maps (Stephanopoulos, 1999).

MFA provides detailed quantification of metabolic fluxes in central metabolism, and was originally applied to the study of microorganisms. As a result of the often unpredictable nature of engineering biological networks, quantification of metabolic fluxes is of utmost importance to understand cellular regulation, identify bottlenecks in product formation, and gain insight to the fundamental processes of biological systems (Boghigian et al., 2010). Beside elucidation of the cellular regulation of the native organism, metabolic flux analysis can be applied to investigate cellular perturbations, such as gene and media manipulation.

In metabolic flux analysis, intracellular metabolite balances are calculated using stoichiometric models for the major intracellular reactions then combined with the fundamental law of mass conservation (Wiechert et al., 2001). Metabolite balances of each metabolite can be written in the form:

Accumulation = [metabolite formation] – [metabolite depletion] (1)

Thus, mass balances around each intracellular metabolite are written to generate a system of linear equations (Varma and Palsson, 1994). As a simple example illustrating the use of a metabolite balance, let us consider the starch synthesis pathway in Figure 1. The following set of coupled Ordinary Differential Equations (ODEs) can be formulated:

$$d\frac{[Glucose\text{-}1\text{-}phosphate]}{dt} = J_8 - J_9 - J_{10} - J_{11} \quad (2)$$

$$d\frac{[glucose]}{dt} = J_1 - J_3 \quad (3)$$

This can be presented in vector matrix notation,

$$\frac{d\overline{\mathbf{X}}}{dt} = \overline{\overline{\mathbf{A}}} * \overline{\mathbf{J}} \tag{4}$$

where, $\overline{\mathbf{X}}$ represents the concentration of the metabolite accumulate in the system; $\overline{\overline{\mathbf{A}}}$ is the stoichiometric matrix, and $\overline{\mathbf{J}}$ is the vector containing the fluxes or reactions. The assumption for the model is that the intracellular metabolite is at pseudo-steady-state, which implies that intracellular pool sizes do not change over the time span during which the experiment takes place. This assumption is justified because of the high enzyme turnover in biochemical systems (Stephanopoulos et al., 1998). As a result, the accumulation rates of the metabolite can be set to zero, as shown below.

$$\frac{d\overline{\mathbf{X}}}{dt} = \overline{\overline{\mathbf{A}}} * \overline{\mathbf{J}} = 0 \tag{5}$$

If the number of reactions is equal to the number of intracellular metabolites, then it is an exactly determined system. Most of the time, however, the number of fluxes is greater than the number of intracellular metabolites. In this case, the solution to the metabolite balancing equations is to have an infinite number of possible solutions. This is referred to as an “under-determined system,” which means that additional experimental measurements, such as extracellular metabolites, are required to reduce the number of unknown fluxes. For large systems, flux analysis becomes more difficult as the number of measurements increase, allowing for the application of linear programming to solve the system of matrix equations. In the case of an “over-determined system,” extra measurements can be used to check the validity of the model.

The limitations of the MFA stoichiometric method result from parallel metabolic pathways, metabolic cycles, reversible or bidirectional reactions, unbalanced enzyme cofactors, and compartmentalization pathways (Stephanopoulos, 1999; Wiechert *et al.*, 2001). These limitations can be monitored by supplementing the stoichiometric MFA with data from ^{13}C-labeling experiment (CLE). CLE is the flux ratio at branch points of the network, reflected in the ^{13}C-labeling pattern of the metabolite, and can provide an additional constraint for the stoichiometric equations to compensate for any lack of measurements (Stephanopoulos, 2002). Many studies have been reported on ^{13}C-labeling experiments that focused on feeding a specifically labeled carbon substrate to the organism of interest, then quantifying the ^{13}C enrichment or isotopomer (isotope isomer) distribution in metabolite carbon atom (Stephanopoulos, 2002). One well known method involves feeding the organism a mixture of uniformly ^{13}C-labeled (U-^{13}C) substrate and a naturally abundant substrate (Szyperski, 1995). When a mixture of uniformly labeled and unlabeled substrates is used, carbon-to-carbon atoms are traced as opposed to fractional enrichments, which allows for lower percentage requirements of uniformly labeled substrate (approximately 10% lower) to achieve adequate labeling data (Szyperski, 1995). By using this method, the metabolic flux information is the extent of scalar coupling between adjacent carbon atoms, as measured by Nuclear Magnetic Resonance (NMR).

Nuclear Magnetic Resonance (NMR) has become one of the key technologies in the elucidation of biosynthetic pathways and metabolite fluxes via quantitative assessment of relative abundances of isotope isomers (isotopomers). NMR is a powerful analytical tool that utilizes an isotopically labeled (e.g. ^{13}C) tracer

compound to identify the distribution of carbon fluxes in complex metabolic networks, specifically providing insight into the labeling patterns of metabolites (Wiechert, 2001). NMR contributes two valuable pieces of flux information: carbon enrichments (the proportion of ^{13}C label that is integrated in an individual carbon position of a metabolite), and isotope isomers or isotopomers (various ^{13}C-^{12}C patterns that can be formed for the same metabolite) (Massou et al., 2007). The isotopomer distribution is determined from the analysis of ^{13}C-^{13}C coupling patterns in labeled metabolite(s) in 1-D ^{13}C or 2-D ^{1}H-^{13}C NMR spectra (Szyperski, 1995). The advantage of 2-D experiments over 1-D experiments is that individual compounds (such as amino acids and glucosyl units in extracted protein) do not need to be separated before the measurement takes place; the two dimensions perform the separation for adequate resolution. Therefore, in this chapter, we focus on the application of 2-D [^{1}H, ^{13}C] Heteronuclear Single Quantum Correlation spectroscopy (HSQC) NMR experiment as a means for providing isotopomer distributions needed for metabolic flux analysis, using soybean embryos cultured *in vitro* as the model plant system.

The term isotopomer is derived from a combination of the terms ‘isotope’ and ‘isomers’, as isotopomers represent the various carbon labeling patterns of a specific metabolite. The isotopomer theory and ^{13}C-NMR application in plant metabolic engineering have been discussed in detail by many research groups around the globe (Allen *et al.*, 2009; Allen *et al.*, 2007; Alonso *et al.*, 2010; Boyle and Morgan, 2009; Dieuaide-Noubhani *et al.*, 2007; Iyer *et al.*, 2007; Ratcliffe and Shachar-Hill, 2001; Schwender, 2008; Shachar-Hill, 2002; Sriram *et al.*, 2004; Stephanopoulos and

Vallino, 1991; Wiechert, 2001). In recent years, the application of carbon labeled metabolic flux analysis (^{13}C-MFA) as a practical tool for generating metabolic flux maps of central metabolic pathways in plants has increased. Some of the plant metabolic flux maps generated for different plant metabolic phenotypes, transgenics, and environmental conditions are listed here: (Allen *et al.*, 2009; Allen and Young, 2013; Alonso *et al.*, 2007; Alonso *et al.*, 2010; Alonso *et al.*, 2007; Alonso *et al.*, 2011; Boyle and Morgan, 2009; Hay and Schwender, 2011, 2011; Iyer *et al.*, 2008; Kruger and Ratcliffe, 2009; Libourel and Shachar-Hill, 2008; Lonien and Schwender, 2009; Morgan *et al.*, 2000; Paula Alonso *et al.*, 2010; Ratcliffe and Shachar-Hill, 2006; Schwender, 2008; Sriram *et al.*, 2004; Sriram *et al.*, 2007).

2D [^{1}H, ^{13}C] HSQC utilizes detailed information about the coupling of the carbon atoms by detecting labeling patterns of target carbon atoms, as well as the adjacent carbon atoms. In addition to the labeling pattern of carbon atoms, 2D HSQC analysis can identify the protons that are attached to either ^{12}C and ^{13}C, which provides detailed information about neighboring carbon atoms in a molecule (Szyperski, 1995). The NMR spectrum can also provide information on the sub-cellular compartmentalization of metabolites (Last *et al.*, 2007). In recent studies, HSQC analysis has been used to distinguish between the glucosyl units on the protein and starch hydrolysates from the soybean tissue, which provides crucial information about different pools of glucose-6-phosphate in the cytosolic and plastidic compartments within plant cells (Sriram *et al.*, 2004; Sriram *et al.*, 2007). This important information about compartmentalization in plants was modeled by the *Arabidopsis* cell (Lonien and Schwender, 2009; Masakapalli *et al.*, 2010; Williams *et al.*, 2008),

and concluded that starch is exclusively synthesized from the plastidic glucose 6-phosphate pool. The peak structure obtained from the HSQC experiment showed multiplet patterns proportional to the isotopomer abundances, and these abundances can be translated to flux information.

The 2D NMR spectrum is useful for plant metabolic flux analysis because it contributes summarized information about the quantities and identities of metabolites present *in vivo.* ^{13}C labeling experiment and 2D NMR application in plant metabolic engineering have been well examined in comprehensive reviews by many researchers over the last decade: (Kruger *et al.*, 2012; Libourel and Shachar-Hill, 2008; Ratcliffe and Shachar-Hill, 2006; Schwender, 2008; Stephanopoulos and Vallino, 1991; Wiechert, 2001). In recent years, elevated interest in omics-based research and plant models to improve crop production for food and feed industries has found the application of carbon labeled metabolic flux analysis (^{13}C-MFA) to be a practical tool for systems biology. Using ^{13}C-MFA, metabolic flux maps were generated for central metabolic pathways with sub-cellular compartmentalization to provide a detailed picture of *in vivo* fluxes in plants. Metabolic flux maps of soybean (Allen *et al.*, 2009; Iyer *et al.*, 2008; Sriram *et al.*, 2004), maize (Alonso *et al.*, 2010; Alonso *et al.*, 2011), sunflower (Alonso *et al.*, 2007), and rapeseed oil (Hay and Schwender, 2011; Schwender *et al.*, 2003; Schwender *et al.*, 2006) are a few of crop metabolic flux map studies. Hence, we are using soybean embryos cultured *in vitro* as the model plant system for the application of 2D [^{1}H,^{13}C] HSQC NMR analysis, with a focus on proteinogenic amino acid and starch labeling.

^{13}C-MFA in Plant Systems

Metabolic networks of plants are often more complex than other organisms. This is due to several aspects of plant life, such as sessile and autotrophic behavior, and having a high degree of sub-cellular compartmentalization (Allen *et al.*, 2009). In addition, metabolite pathways of plants are separated into compartments such as cytosol, plastid, and mitochondrion, which are interconnected by transporters. This complicates the quest for a detailed and predictive understanding of the regulation of plant metabolism (Carrari *et al.*, 2003; Schwender, 2008). Therefore, it remains difficult to manipulate carbon flows through the central metabolic network of plants in a predictable way (Kruger *et al.*, 2012; Kruger *et al.*, 2008; Sweetlove *et al.*, 2008).

The application of a pseudo-steady-state carbon labeled metabolic flux analysis (^{13}C-MFA) as a practical tool for generating metabolic flux maps of central metabolic pathways in plants provides new opportunities for analyzing plant metabolic phenotypes and environmental conditions (Allen *et al.*, 2009; Alonso *et al.*, 2010; Iyer *et al.*, 2008; Kruger and Ratcliffe, 2009; Libourel and Shachar-Hill, 2008; Lonien and Schwender, 2009; Masakapalli *et al.*, 2009; O'Grady *et al.*, 2012; Ratcliffe and Shachar-Hill, 2006; Schwender, 2008; Sriram *et al.*, 2004; Sriram *et al.*, 2007). As mentioned before, this approach requires the formulation of balanced equations around each metabolite in the network. These metabolite balances require extra-cellular measurements and intracellular measurements. The labeling studies are performed under metabolic and isotopic pseudo-steady-state conditions. Sub-cellular compartmentalization of pathways can be incorporated into a model that describes the redistribution of the label and flux maps of central carbon metabolism in plant cells to

attempt to distinguish between fluxes in the cytosol, mitochondria, and plastids (Sriram *et al.*, 2004; Sriram *et al.*, 2007).

A mathematical framework (NMR2Flux) for metabolic flux analysis has been developed in our lab, and has been used to successfully explain various aspects of central carbon metabolism in developing soybean embryos (Iyer *et al.*, 2008; Sriram *et al.*, 2004). Metabolic flux analysis utilizes the patterns of labeled metabolites resulting from isotopic labeling experiments, as well as uptake and output measurements, to establish internal metabolites flux values. The metabolite flux map then provided insights into the carbon conversion processes of embryo metabolites. For example, Sriram et al., (2004) suggested a small flux from malate into pyruvate in the plastid compartment, indicating that malate provides little carbon for fatty acid biosynthesis. In addition, according to Iyer et al., (2008), fluxes through cytosolic phospheonolpuruvate carboxylase (*ppc*), plastidic malic enzyme (me^P), and the malate transporters varied considerably with temperature during the growth and culture treatment. The me^P and *ppc* enzymes and malate transporters (*malT1* and *malT2*) contribute to the pyruvate pool in the plastid, which is the metabolic branch-point of protein versus oil formation in soybean embryos. Hence, these reactions may be potential targets for improving the current understanding of the regulation of carbon flow into protein and lipid in soybeans. In this research, we apply NMR2Flux (with a few minor changes in metabolic network topology model that was verified from soybean zygotic embryos transcriptomic profiling and extracellular measurements) to determine whether SHaM embryos are suitable models for studying the mechanisms that control assimilate partitioning in seed.

Literature Review: Soybean

The inverse relationship between protein and oil content has resulted from correlations found in field studies. Studies from the early 1970s to 2000s have identified the correlation between soybean seed protein, oil, and oligosaccharides. Hymowitz et al., 1972 observed a positive relationship between oligosaccharides (sucrose, raffinose, and stachyose) and protein ($R^2 = 0.41$) in seed from 60 soybean plants [*Glycine max* (L.) Merrill] with varying protein and oil contents. Wilcox and Shibles (2001) investigated the relationships in soybean seed between protein, oil, oligosaccharides, and sulfur among random progenies of a cross between parents differing in seed oil and protein content. Carbohydrate concentrations were not related to seed yield, though protein yield increased at the expensive of oil, total carbohydrate, and sucrose(Wilcox and Shibles, 2001). The results of these field studies indicate a complex relationship between oligosaccharides, protein, and oil within the different soybean lines of high and low protein and oil content.

In whole-plant experiments from the field, it is difficult to differentiate seed responses from environment influences, and the impact of the mother plant on protein and oil content. *In vitro* growth of soybean seeds has been developed as a method to study soybean seed composition, with the ability to control interactions with the mother plant, as well as environment factors such as photoperiod, light intensity, temperature, and soil fertility (Obendorf et al., 1983; Thompson et al., 1977). Soybean embryos can be grown *in vitro* using defined media containing sucrose (as a main carbon source), one amino acid (as a reduced nitrogen source), and minerals.

Experiments from *in vitro* cultured soybean zygotic embryos, at varying glutamine and sucrose concentrations, suggest that there is a weak relationship between protein and oil content (Pipolo *et al.*, 2004; Saravitz and Raper, 1995). *In vitro* studies show that the partitioning of protein and oil can be manipulated through modification of nutrient media composition. Saravitz and Raper (1995) evaluated the influence of sucrose (1.5, 15, and 150 mM) and glutamine (0.6, 6, and 120 mM) concentrations on soybean seed protein, oil, and carbohydrates (estimated) contents. Their results indicate that oil content increased as the carbon-to-nitrogen (C: N) mole ratios increased, to some extent. However, when combining all of the sucrose and glutamine treatments, the result indicates a weakly inverse relationship between protein and oil content, with a linear coefficient of R^2=0.42. Pipolo et al. (2004) studied soybean embryos grown in media conditions containing glutamine concentrations ranging from 20 to 80 mM at a sucrose concentration of 204.5 mM; and, once again, the protein and oil were inversely correlated (R^2 = 0.44).

In recent studies of *in vitro* zygotic embryos with a range of C: N mole ratios from 6 to 91 (10 media conditions), with glucose and sucrose as a main carbon sources and two amino acids of glutamine and asparagines (Allen and Young, 2013). The data indicated that the protein content of embryos increased as the C: N ratios decreased (14 to 47 % Dry weight). Examining the results from the report illustrate that the oil content was not perturbed, but the starch content displayed a trend of increased as the C: N ratios increased. Three conditions of C: N ratios of 13, 21 and 37 were chosen for furthermore investigation with labeling glutamine, glucose, and asparagines. The result demonstrates that the role of glutamine as an reduced nitrogen

source in the media dictated the protein biosynthesis. The glutamine labeling experiment indicated that approximately 40% of carbon to amino acids and nearly 15% of carbon to fatty acid synthesis was supplied mainly from glutamine. This illustrates the difficulty in identifying the controlling point of carbon partitioning in protein and oil synthesis with only the manipulation of the carbon and nitrogen sources treatment to the soybean embryos. Our study aims to elucidate the carbon partitioning of protein and oil with combination treatment of genetic and media manipulation to soybean embryos.

There are many advantages to using somatic embryos as a model to understand the metabolism of developing zygotic (seed) embryos. Soybean somatic embryos, particularly those grown in Soybean Histodifferentiation and Maturation media (SHaM) (Schmidt et al., 2005), accumulate seed-specific protein and triglycerides, with compositions comparable to soybean seed. A recent study has found that somatic embryos accumulate seed-specific storage proteins, such as β-conglycinin and glycinin, in a manner similar to that of seed development (Nishizawaa and Ishimoto, 2009). In addition to the accumulation of seed specific proteins, SHaM embryos, like seeds, have more total triacylglycerols (TAG) in the total lipid extract than other soybean somatic embryos (Schmidt et al., 2005).

Most importantly, soybean somatic embryos in SHaM media can be used as a tool to study the accumulation of carbon into storage reserves in wild-type or transgenic cultures under different environmental perturbations, for example, under varying of carbon-to-nitrogen (C: N) mole ratios in the media. Additionally, soybean somatic embryos can be tested more quickly than zygotic embryos for genotype-phenotype

relationships in a transgenic research program. Somatic embryos in SHaM media can be recovered into whole plants through controlled growth in maturation and germination medium (Bailey et al., 1993). Furthermore, the cycle time from transformation-to-harvest of mature somatic embryos in SHaM media is approximately 8 to 10 weeks, at which time stable transgenic events can be identified and maintained for experimentation. In contrast, it takes approximately 11 months to generate seed bearing transgenic plants, and several more generations are required to obtain homozygous plants capable for generating zygotic embryos for study. The use of somatic embryos in SHaM media, therefore, makes it feasible to transgenically manipulate metabolic control points as a way to test hypotheses generated by metabolic flux analysis studies.

Application of genetic engineering to soybean somatic embryos can be used to enhance the understanding of protein and lipid relationship. Direct manipulation of a single enzymatic reaction, such as phosphoglucomutase (PGM), in the plastidic compartment of soybean embryos can suppress the accumulation of transient starch. PGM is a widely distributed enzyme that catalyzes the readily reversible interconversion of glucose-1-phosphate (Glc-1-P) and glucose-6-phosphate (Glc-6-P). There are two PGM isoforms in plants, one localized in the plastids and the other in the cytosol (Schnarrenberger et al., 1973). The cytosolic isoform is involved in sucrose catabolism to provide intermediates for glycolysis and substrates for the syntheses of many cellular constituents (Manjunath et al., 1998). The plastidic PGM is essential for photosynthesis starch in the leaves (Dietz, 1987), and also plays an

essential role in the degradation of assimilated starch from glucose nutrition in the plant (Hattenbach and Heineke, 1999).

There are transgenic soybean lines in which the plastidic PGM has been severely down regulated, ultimately making the seed incapable of producing starch during development (Allen et al., 2008). In the PGM-KO line, transformation is stable, and somatic embryos express a low starch phenotype, while PGM-null lines undergo the same method of transformation, while preventing expression of the same low starch phenotype. Somatic embryos that are generated from these stably transformed lines, PGM-KO and PGM-null, can be maintained in culture and used for experimentation. The PGM-null line is important for two reasons: first, because the PGM-KO mutation is created by transgenic means, it is necessary to have a control that has undergone the same process (transformation, selection, differentiation, and plant regeneration), but does not express the low starch phenotype. Second, the PGM transgenic events are created in Pioneer Hi-Bred soybean variety 93B92, and not cultivar *Jack*. It is therefore important to have a control line that has the same base genetics against which comparisons can be made. The PGM-KO soybean somatic embryos in SHaM culture have a starch content of approximately 0.07 ± 0.02 %, on dry weight basis, compared with 7.10 ± 1.21 % in PGM-null, the control variable (data were provided by Pioneer Hi-Bred). The plastidic PGM not only plays an essential role in starch synthesis, but also has a significant impact on the depositing of other storage products in soybean somatic embryos and seeds (Periappuram et al., 2000). Therefore, studying soybean somatic embryos of PGM-KO and PGM-null in SHaM media under

different carbon and nitrogen ratios may help decipher the interaction between protein and oil synthesis.

Storage metabolism in plants involves different cellular compartments such as the cytosol, plastids, and mitochondria, with interacting pathways between compartments (Bowsher and Tobin, 2001; Schwender et al., 2003). The metabolic pathways of the major storage compounds of protein, lipid, and starch are well-described (Borisjuk et al., 2005; Litterer et al., 2006), but the functioning of these pathways, which determine carbon partitioning into major storage compounds, remains largely unknown. Protein accumulation in seeds depends mainly on nitrogen uptake and availability (Golombek *et al.*, 2001; Miranda *et al.*, 2001; Salon *et al.*, 2001), and glutamine is the preferred source of nitrogen for developing soybean seeds grown *in vitro.* Developing cotyledons have a limited capacity for the assimilation of other forms of nitrogen, and essentially no capacity for the assimilation of inorganic nitrogen (Haga and Sodek, 1987; Thompson *et al.*, 1977). Conversion of glutamine to the other amino acids required for storage protein synthesis demands that carbon skeletons derived from the metabolism of sucrose be imported from the phloem (Miflin and Lea, 1977). In the case of glutamine, its amide-N group must be donated to a suitable C-skeleton acceptor molecule, such as 2-Oxoglutarate (a product of TCA cycle in the mitochondria), which can then be utilized further in the formation of other amino acids (Weigelt et al., 2008). Lipid synthesis utilizes metabolites derived from sucrose catabolism (Hernandez-Sebastia et al., 2005), and thus, storage protein synthesis and lipid synthesis both require C-skeletons derived from (Hernandez-Sebastia *et al.*, 2005)imported sucrose during seed development. This information

suggests that the inverse relationship between seed storage protein and lipid concentration might be related to the regulation of carbon flux between these competing synthetic pathways, and the interaction between cellular compartments. A better understanding of how carbon flux between protein and lipid synthesis is regulated might lead to the development of molecular strategies to improve soybean seed quality.

Organization of Thesis

This chapter introduced the theoretical background and methodology of metabolic flux analysis, and its importance in understanding the preferred direction of carbon flow in plant systems, particularly in soybean embryos. Furthermore, this chapter presents a literature review of current thoughts on varying carbon to nitrogen ratios in different metabolites, compartmental and metabolic flux analysis plant systems, the innovation of MFA in plants, and the need for an understanding of carbon regulation in the central metabolites of compartments of soybean somatic embryos wild-type and phosphoglucomutase (PGM) mutant in SHaM media.

Chapter 2 reports the response of soybean somatic to the influence of different carbon to nitrogen ratios in SHaM media. The results of this experiment provide data for identifying which media conditions will be used for carbon labeled metabolite flux analysis in an effort to understand the partitioning of protein and oil in soybean somatic embryos.

Chapter 3 describes in detail the application of the flux analysis tool NMR2Flux (Sriram *et al.*, 2004) to determine whether soybean somatic embryos are suitable

models for studying the mechanisms controlling assimilate partitioning in seed. Please notice that this chapter contains confidential information.

Chapter 4 presents the metabolic flux maps with a comparison of PGM mutant embryos with the control embryos, along with the impact of sucrose level in the control and PGM embryos. Please notice that this chapter contains confidential information.

Chapter 5 summarizes the conclusions of this work and provides future direction.

References

Allen DK, Libourel IGL, Shachar-Hill Y. 2009. Metabolic flux analysis in plants: coping with complexity. *Plant Cell and Environment* 32, 1241-1257.

Allen DK, Ohlrogge JB, Shachar-Hill Y. 2009. The role of light in soybean seed filling metabolism. *Plant Journal* 58, 220-234.

Allen DK, Shachar-Hill Y, Ohlrogge JB. 2007. Compartment-specific labeling information in C-13 metabolic flux analysis of plants. *Phytochemistry* 68, 2197-2210.

Allen DK, Young JD. 2013. Carbon and Nitrogen Provisions Alter the Metabolic Flux in Developing Soybean Embryos. *Plant Physiology* 161, 1458-1475.

Allen SM, Butler KH, Carlson TJ, Hitz WD, Stoop JM. 2008. Plastidic phosphoglucomutase genes. Vol. US 20080066204A1: EI du Pont de Nemours and Company.

Alonso AP, Dale VL, Shachar-Hill Y. 2010. Understanding fatty acid synthesis in developing maize embryos using metabolic flux analysis. *Metabolic Engineering* 12, 488-497.

Alonso AP, Goffman FD, Ohlrogge JB, Shachar-Hill Y. 2007. Carbon conversion efficiency and central metabolic fluxes in developing sunflower (Helianthus annuus L.) embryos. *Plant Journal* 52, 296-308.

Alonso AP, Piasecki RJ, Wang Y, LaClair RW, Shachar-Hill Y. 2010. Quantifying the Labeling and the Levels of Plant Cell Wall Precursors Using Ion Chromatography Tandem Mass Spectrometry. *Plant Physiology* 153, 915-924.

Alonso AP, Raymond P, Rolin D, Dieuaide-Noubhani M. 2007. Substrate cycles in the central metabolism of maize root tips under hypoxia. *Phytochemistry* 68, 2222-2231.

Alonso AP, Val DL, Shachar-Hill Y. 2011. Central metabolic fluxes in the endosperm of developing maize seeds and their implications for metabolic engineering. *Metabolic Engineering* 13, 96-107.

Bailey MA, Boerma HR, Parrott WA. 1993. Genotype-Specific Optimization of Plant-Regeneration from Somatic Embryos of Soybean. *Plant Science* 93, 117-120.

Boghigian BA, Seth G, Kiss R, Pfeifer BA. 2010. Metabolic flux analysis and pharmaceutical production. *Metabolic Engineering* 12, 81-95.

Borisjuk L, Nguyen TH, Neuberger T, et al. 2005. Gradients of lipid storage, photosynthesis and plastid differentiation in developing soybean seeds. *New Phytologist* 167, 761-776.

Bowsher CG, Tobin AK. 2001. Compartmentation of metabolism within mitochondria and plastids. *Journal of Experimental Botany* 52, 513-527.

Boyle NR, Morgan JA. 2009. Flux balance analysis of primary metabolism in Chlamydomonas reinhardtii. *BMC Systems Biology* 3, (7 January 2009).

Carrari F, Urbanczyk-Wochniak E, Willmitzer L, Fernie AR. 2003. Engineering central metabolism in crop species: learning the system. *Metabolic Engineering* 5, 191-200.

Clemente TE, Cahoon EB. 2009. Soybean Oil: Genetic Approaches for Modification of Functionality and Total Content. *Plant Physiology* 151, 1030-1040.

Dietz K. 1987. Control function of hexosemonophosphate isomerase and phosphoglucomutase in starch synthesis in leaves. *In J Biggins, ed, Proceedings of the 7th International Congress on Photosynthesis* 3, 329-332.

Dieuaide-Noubhani M, Alonso AP, Rolin D, Eisenreich W, Raymond P. 2007. *Metabolic flux analysis: recent advances in carbon metabolism in plants*.

Durot M, Bourguignon PY, Schachter V. 2009. Genome-scale models of bacterial metabolism: reconstruction and applications. *FEMS Microbiology Reviews* 33, 164-190.

Feist AM, Herrgard MJ, Thiele I, Reed JL, Palsson BO. 2009. Reconstruction of biochemical networks in microorganisms. *Nature Reviews Microbiology* 7, 129-143.

Golombek S, Rolletschek H, Wobus U, Weber H. 2001. Control of storage protein accumulation during legume seed development. *Journal of Plant Physiology* 158, 457-464.

Haga KI, Sodek L. 1987. Utilization of Nitrogen-Sources by Immature Soybean Cotyledons in Culture. *Annals of Botany* 59, 597-601.

Hartwig EE, Kilen TC. 1991. Yield and Composition of Soybean Seed from Parents with Different Protein, Similar Yield. *Crop Science* 31, 290-292.

Hattenbach A, Heineke D. 1999. On the role of chloroplastic phosphoglucomutase in the regulation of starch turnover. *Planta* 207, 527-532.

Hay J, Schwender J. 2011. Computational analysis of storage synthesis in developing Brassica napus L. (oilseed rape) embryos: flux variability analysis in relation to (13)C metabolic flux analysis. *Plant Journal* 67, 513-525.

Hay J, Schwender J. 2011. Metabolic network reconstruction and flux variability analysis of storage synthesis in developing oilseed rape (Brassica napus L.) embryos. *Plant Journal* 67, 526-541.

Hernandez-Sebastia C, Marsolais F, Saravitz C, Israel D, Dewey RE, Huber SC. 2005. Free amino acid profiles suggest a possible role for asparagine in the control of storage-product accumulation in developing seeds of low- and high-protein soybean lines. *Journal of Experimental Botany* 56, 1951-1963.

Iyer VV, Sriram G, Fulton DB, Zhou R, Westgate ME, Shanks JV. 2008. Metabolic flux maps comparing the effect of temperature on protein and oil biosynthesis in developing soybean cotyledons. *Plant Cell and Environment* 31, 506-517.

Iyer VV, Sriram G, Shanks JV. 2007. Metabolic flux maps of central carbon metabolism in plant systems - Isotope labeling analysis. *Concepts in Plant Metabolomics*, 125-144.

Kell DB. 2004. Metabolomics and systems biology: making sense of the soup. *Current Opinion in Microbiology* 7, 296-307.

Kruger NJ, Masakapalli SK, Ratcliffe RG. 2012. Strategies for investigating the plant metabolic network with steady-state metabolic flux analysis: lessons from an Arabidopsis cell culture and other systems. *Journal of Experimental Botany* 63, 2309-2323.

Kruger NJ, Ratcliffe RG. 2009. Insights into plant metabolic networks from steady-state metabolic flux analysis. *Biochimie* 91, 697-702.

Kruger NJ, Troncoso-Ponce MA, Ratcliffe RG. 2008. 1H NMR metabolite fingerprinting and metabolomic analysis of perchloric acid extracts from plant tissues. *Nat. Protocols* 3, 1001-1012.

Last RL, Jones AD, Shachar-Hill Y. 2007. Towards the plant metabolome and beyond. *Nature Reviews Molecular Cell Biology* 8, 167-174.

Libourel IGL, Shachar-Hill Y. 2008. Metabolic flux analysis in plants: from intelligent design to rational engineering. *Annual Review of Plant Biology* 59, 625-650.

Litterer LA, Plaisance KL, Schnurr JA, Storey KK, Jung HJG, Gronwald JW, Somers DA. 2006. Biosynthesis of UDP-glucuronic acid in developing soybean embryos: possible role of UDP-sugar pyrophosphorylase. *Physiologia Plantarum* 128, 200-211.

Lonien J, Schwender J. 2009. Analysis of Metabolic Flux Phenotypes for Two Arabidopsis Mutants with Severe Impairment in Seed Storage Lipid Synthesis. *Plant Physiology* 151, 1617-1634.

Lu C, Napier JA, Clemente TE, Cahoon EB. 2011. New frontiers in oilseed biotechnology: meeting the global demand for vegetable oils for food, feed, biofuel, and industrial applications. *Current Opinion in Biotechnology* 22, 252-259.

Manjunath S, Lee CHK, VanWinkle P, Bailey-Serres J. 1998. Molecular and biochemical characterization of cytosolic phosphoglucomutase in maize - Expression during development and in response to oxygen deprivation. *Plant Physiology* 117, 997-1006.

Masakapalli SK, Lay PL, Huddleston JE, Pollock NL, Kruger NJ, Ratcliffe RG. 2009. Subcellular flux analysis of central metabolism in a heterotrophic Arabidopsis thaliana cell suspension using steady-state stable isotope labeling. *Plant Physiology*, pp.109.151316.

Masakapalli SK, Le Lay P, Huddleston JE, Pollock NL, Kruger NJ, Ratcliffe RG. 2010. Subcellular Flux Analysis of Central Metabolism in a Heterotrophic Arabidopsis Cell Suspension Using Steady-State Stable Isotope Labeling. *Plant Physiology* 152, 602-619.

Massou S, Nicolas C, Letisse F, Portais JC. 2007. NMR-based fluxomics: Quantitative 2D NMR methods for isotopomers analysis. *Phytochemistry* 68, 2330-2340.

Medini D, Serruto D, Parkhill J, Relman DA, Donati C, Moxon R, Falkow S, Rappuoli R. 2008. Microbiology in the post-genomic era. *Nature Reviews Microbiology* 6, 419-430.

Miflin BJ, Lea PJ. 1977. Amino- Acid Metabolism *Briggs, Winslow R. (Ed.). Annual Review of Plant Physiology, Vol. 28. X+615p. Illus. Annual Reviews Inc.: Palo Alto, Calif., USA. Isbn 0-8243-0628-7*, 299-329.

Miranda M, Borisjuk L, Tewes A, Heim U, Sauer N, Wobus U, Weber H. 2001. Amino acid permeases in developing seeds of Vicia faba L.: expression precedes storage protein synthesis and is regulated by amino acid supply. *Plant Journal* 28, 61-71.

Morgan JA, Barney CS, Penn AH, Shanks JV. 2000. Effects of buffered media upon growth and alkaloid production of Catharanthus roseus hairy roots. *Applied Microbiology and Biotechnology* 53, 262-265.

Nielsen J, Oliver S. 2005. The next wave in metabolome analysis. *Trends in Biotechnology* 23, 544-546.

Nishizawaa K, Ishimoto M. 2009. Maturation of somatic embryos as a model for soybean seed development. *Plant Biotechnology* 26, 543-550.

O'Grady J, Schwender J, Shachar-Hill Y, Morgan JA. 2012. Metabolic cartography: experimental quantification of metabolic fluxes from isotopic labelling studies. *Journal of Experimental Botany* 63, 2293-2308.

Obendorf RL, Rytko GT, Byrne MC. 1983. Soya Bean Seed Growth and Maturation by In Vitro Pod Culture. *Ann Bot* 51, 217-227.

Paula Alonso A, Dale VL, Shachar-Hill Y. 2010. Understanding fatty acid synthesis in developing maize embryos using metabolic flux analysis. *Metabolic Engineering* 12, 488-497.

Periappuram C, Steinhauer L, Barton DL, Taylor DC, Chatson B, Zou J. 2000. The plastidic phosphoglucomutase from Arabidopsis. A reversible enzyme reaction with an important role in metabolic control. *Plant Physiology* 123, 1197-1197.

Pipolo AE, Sinclair TR, Camara GMS. 2004. Protein and oil concentration of soybean seed cultured in vitro using nutrient solutions of differing glutamine concentration. *Annals of Applied Biology* 144, 223-227.

Ratcliffe RG, Shachar-Hill Y. 2001. Probing plant metabolism with NMR. *Annual Review of Plant Physiology and Plant Molecular Biology* 52, 499-526.

Ratcliffe RG, Shachar-Hill Y. 2006. Measuring multiple fluxes through plant metabolic networks. *Plant Journal* 45, 490-511.

Salon C, Munier-Jolain NG, Duc G, Voisin AS, Grandgirard D, Larmure A, Emery RJN, Ney B. 2001. Grain legume seed filling in relation to nitrogen acquisition: A review and prospects with particular reference to pea. *Agronomie* 21, 539-552.

Saravitz CH, Raper CD. 1995. Responses to Sucrose and Glutamine by Soybean Embryos Growth *In-Vitro*. *Physiologia Plantarum* 93, 799-805.

Schmidt MA, Tucker DM, Cahoon EB, Parrott WA. 2005. Towards normalization of soybean somatic embryo maturation. *Plant Cell Reports* 24, 383-391.

Schnarrenberger C, Oeser A, Tolbert NE. 1973. Two isoenzymes each of glucose-6-phosphate dehydrogenase and 6-phosphogluconate dehydrogenase in spinach leaves. *Archives of Biochemistry and Biophysics* 154, 438-448.

Schwender J. 2008. Metabolic flux analysis as a tool in metabolic engineering of plants. *Current Opinion in Biotechnology* 19, 131-137.

Schwender J, Ohlrogge JB, Shachar-Hill Y. 2003. A flux model of glycolysis and the oxidative pentosephosphate pathway in developing Brassica napus embryos. *Journal of Biological Chemistry* 278, 29442-29453.

Schwender J, Shachar-Hill Y, Ohlrogge JB. 2006. Mitochondrial metabolism in developing embryos of Brassica napus. *Journal of Biological Chemistry* 281, 34040-34047.

Shachar-Hill Y. 2002. Nuclear magnetic resonance and plant metabolic engineering. *Metabolic Engineering* 4, 90-97.

Sinclair TR, de Wit CT. 1975. Photosynthate and Nitrogen Requirements for Seed Production by Various Crops. *Science* 189, 565-567.

Sriram G, Fulton DB, Iyer VV, Peterson JM, Zhou R, Westgate ME, Spalding MH, Shanks JV. 2004. Quantification of compartmented metabolic fluxes in developing soybean embryos by employing biosynthetically directed fractional ^{13}C labeling, two-dimensional [^{13}C, ^{1}H] nuclear magnetic resonance, and comprehensive isotopomer balancing. *Plant Physiology* 136, 3043-3057.

Sriram G, Fulton DB, Iyer VV, Peterson JM, Zhou RL, Westgate ME, Spalding MH, Shanks JV. 2004. Quantification of compartmented metabolic fluxes in developing soybean embryos by employing Biosynthetic ally directed fractional C-13 labeling, C-13, H-1 two-dimensional nuclear magnetic resonance, and comprehensive isotopomer balancing. *Plant Physiology* 136, 3043-3057.

Sriram G, Fulton DB, Shanks JV. 2007. Flux quantification in central carbon metabolism of Catharanthus roseus hairy roots by C-13 labeling and comprehensive bondomer balancing. *Phytochemistry* 68, 2243-2257.

Sriram G, Iyer VV, Fulton DB, Shanks JV. 2007. Identification of hexose hydrolysis products in metabolic flux analytes: A case study of levulinic acid in plant protein hydrolysate. *Metabolic Engineering* 9, 442-451.

Stephanopoulos G. 1999. Metabolic Fluxes and Metabolic Engineering. *Metabolic Engineering* 1, 1-11.

Stephanopoulos G. 2002. Metabolic engineering: Perspective of a chemical engineer. *Aiche Journal* 48, 920-926.

Stephanopoulos G, Vallino JJ. 1991. Network Rigidity and Metabolic Engineering in Metabolite Overproduction. *Science* 252, 1675-1681.

Stephanopoulos GN, Aristidou AA, Nielsen J. 1998. Metabolic Flux Analysis. *Metabolic Engineering*. San Diego: Academic Press, 309-351.

Sweetlove LJ, Fell D, Fernie AR. 2008. Getting to grips with the plant metabolic network. *Biochem J* 409, 27-41.

Szyperski T. 1995. Biosynthetically directed fractional ^{13}C-labeling of proteinogenic amino acids. An efficient analytical tool to investigate intermediary metabolism. *European Journal of Biochemistry* 232, 433-448.

Szyperski T. 1995. Biosynthetically Directed Fractional C-13-Labeling of Proteinogenic Amino-Acids - An Efficient Analytical Tool to Investigate Intermediary Metabolism. *European Journal of Biochemistry* 232, 433-448.

Thompson JF, Madison JT, Muenster AME. 1977. *In Vitro* Culture of Immature Cotyledons of Soybean (*Glycine-Max-L-Merr*) *Annals of Botany* 41, 29-39.

van der Werf MJ, Overkamp KM, Muilwijk B, Coulier L, Hankemeier T. 2007. Microbial metabolomics: Toward a platform with full metabolome coverage. *Analytical Biochemistry* 370, 17-25.

Varma A, Palsson BO. 1994. Stoichiometric Flux Balance Models Quantitatively Predict Growth and Metabolic By-Product Secretion in Wild-Type *Escherichia-Coli W3110*. *Applied and Environmental Microbiology* 60, 3724-3731.

Weigelt K, Kuster H, Radchuk R, Muller M, Weichert H, Fait A, Fernie AR, Saalbach I, Weber H. 2008. Increasing amino acid supply in pea embryos reveals specific interactions of N and C metabolism, and highlights the importance of mitochondrial metabolism. *Plant Journal* 55, 909-926.

Wiechert W. 2001. C-13 metabolic flux analysis. *Metabolic Engineering* 3, 195-206.

Wiechert W, Mollney M, Petersen S, de Graaf AA. 2001. A universal framework for C-13 metabolic flux analysis. *Metabolic Engineering* 3, 265-283.

Wilcox JR, Shibles RM. 2001. Interrelationships among seed quality attributes in soybean. *Crop Science* 41, 11-14.

Williams TCR, Miguet L, Masakapalli SK, Kruger NJ, Sweetlove LJ, Ratcliffe RG. 2008. Metabolic network fluxes in heterotrophic arabidopsis cells: stability of the flux distribution under different oxygenation conditions. *Plant Physiology* 148, 704-718.

Figure caption

Figure 1: A simple example of starch synthesis network, from the metabolic network the stoichiometric relations can be immediately derived.

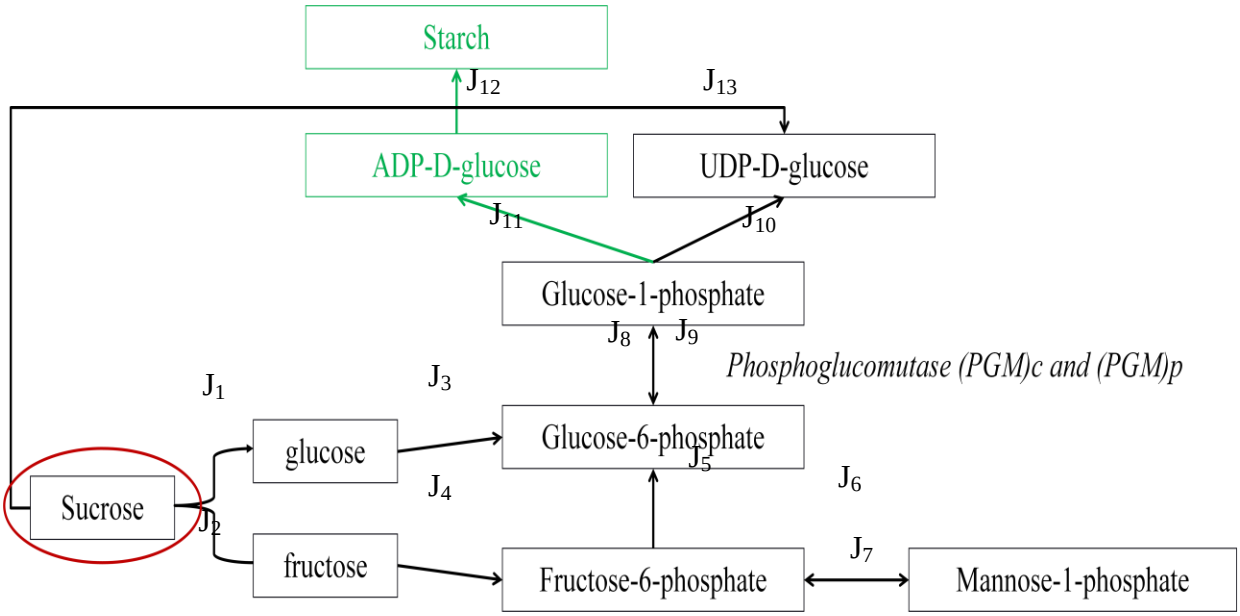


Figure 1

CHAPTER 2

INFLUENCE OF CARBON-TO-NITROGEN RATIOS ON SOYBEAN SOMATIC EMBRYOS (*CV. JACK*) GROWTH AND COMPOSITION

A paper accepted to publish in the *Journal of Experimental Botany*

Quyen Truong[1], Kaelynn Koch[1], Jong Moon Yoon[1], John D. Everard[2], Jacqueline V. Shanks[1]

[1]Department of Chemical and Biological Engineering, Iowa State University, Ames, IA 50010, USA, [2]DuPont Agricultural Biotechnology Research and Development, DuPont Experimental Station, Wilmington, Delaware 19880, USA

Abstract

Soybean [*Glycine max* (L.) Merr.] seed are valued for their protein and oil content. Soybean somatic embryos cultured in Soybean Histodifferentiation and Maturation (SHaM) media were examined for their suitability as a model system for developing an understanding of assimilate partitioning and metabolic control points for protein and oil biosynthesis in soybean seed. This report describes the growth dynamics and compositional changes of SHaM embryos in response to change in the carbon-to-nitrogen ratio of the media. We postulated that at media compositions that were sufficient to support maximal growth rates, changes in the C: N ratio are likely to influence the partitioning of resources between the various storage products, especially protein and oil. As postulated, at steady-state growth rates embryo protein content was strongly correlated to decreasing C: N ratios and increasing glutamine consumption rates. However, oil content remained relatively unchanged across the C: N ratio range tested and resources were instead directed towards the starch and residual biomass (estimated by mass balance) pools in response to increasing C: N

ratios. Protein and oil were inversely related only at media sucrose concentrations below 88 mM, where carbon limited growth and no starch was found to accumulate in the tissues. These observations and the high reproducibility in the data, indicate that SHaM embryos are an ideal model system for the application of metabolic flux analysis studies designed to test hypotheses regarding assimilate partitioning in developing soybean seeds.

1 Introduction

Soybean [*Glycine max* (L.) Merr.] seed are valued for their protein and oil content, and when processed, provide the largest global source of vegetable protein for animal feed. Soybean seeds have a very high protein content (40%), but uniquely they are also high in oil (20%) (Sinclair and de Wit, 1975). The conundrum is that the oil and protein content of soybean seeds are inversely related (Hartwig and Kilen, 1991); specifically a 1% reduction in oil content will lead to a 2% increase in protein content (Clemente and Cahoon, 2009). Although carbon shifts between protein and oil content, the underlying metabolic control mechanisms are not fully understood, and strong metabolic links between oil and storage protein synthesis are not apparent (Schwender *et al.*, 2003). The overall challenge is to increase the oil content of soybeans at the expense of other seed components (starch, soluble sugars, and cell wall material), while maintaining protein content at a level that the market demands; i.e., a dehulled, defatted meal with 47.5 - 49% protein (http://www.soymeal.org/composition). Biotechnology and metabolic engineering hold promise to accomplish this challenging task (Lu *et al.*, 2011), but in order to

make progress we need to fully understand the control points of carbon partitioning which dictate the inverse relationship between protein and oil contents in developing soybean seed.

The genome sequence of soybean is available for the application of functional genomics to characterize the genetic basis of important traits (Schmutz *et al.*, 2010). Metabolic pathways for soybean seed components and their regulation at the molecular level should provide important information for improvement of seed quality by genetic engineering (Nishizawa *et al.*, 2010). However, due to the long life cycle of soybeans, using developing seed (zygotic embryos) as an experimental system is neither an efficient nor practical approach for screening the large number of genes (by knockout, over-expression, or under-expression) implicated in seed quality. Approximately 11 months are required to generate seed bearing transgenic plants, and several more generations are required to obtain homozygous plants capable of generating zygotic embryos for study. One response to the limitation of zygotic embryos to support discovery research programs aimed at modifying seed composition has been the development of soybean somatic embryo model systems. Somatic embryos are the target tissue for one commonly used method of soybean genetic transformation (Finer and McMullen, 1991). For certain transgenic seed-specific traits, such as modified fatty acid profiles, somatic embryos have been shown to provide an excellent "preview" of the ultimate composition of mature soybean seed, within 10 weeks of the initial transformation (Kinney, 1996). Despite the success of somatic embryo systems for the development of transgenic soybeans with modified seed quality, the systems have proven to be much more challenging when

applied to evaluate transgenic approaches aimed at quantitatively increasing the total oil or protein content of soybean seed. Soybean somatic embryos accumulate seed specific storage proteins, such as β-conglycinin and glycinin, and oils in the form of triacylglycerides (TAG) in a manner similar to that observed during seed development (Nishizawa and Ishimoto, 2009; Schmidt *et al.*, 2005). However, heterogeneity in embryo development and composition within cultures result in such extreme variation in oil and protein content such that transgenic approaches to change them are very difficult to detect within the natural variation of the system. A breakthrough occurred with the introduction of a somatic embryo model system based on Soybean Histodifferentiation and Maturation (SHaM) media (Schmidt *et al.*, 2005). SHaM embryos show greater developmental uniformity, have compositions that are more seed like (Schmidt *et al.*, 2005) and have proven to be an excellent system for testing genes that lead to oil content increases in mature seed (Meyer *et al.*, 2012).

SHaM embryos are now a proven system for the evaluation of transgenic approaches to improve soybean quality; however, it has yet to be determined whether they provide a suitable test system for developing an understanding of the biochemical and physiological mechanisms underlying resource partitioning in developing soybean seed. Although studies of assimilate partitioning between storage products under different C: N ratios have been performed in zygotic soybean embryos (Allen and Young, 2013; Hayati *et al.*, 1996; Pipolo *et al.*, 2004; Saravitz and Raper, 1995), such studies have not been performed with somatic embryos. In this report we describe an in depth study of the growth dynamics and compositional changes of

SHaM embryos in response to changes in the carbon to nitrogen ratio of the media, with two goals in mind. The first goal was to test a working hypothesis that at constant growth rates, changes in the relative proportions of sucrose (carbon source) and glutamine (nitrogen source) supply to the developing tissue will influence partitioning of resources between the protein and oil pools. The second goal was to identify which sets of C: N ratios in the SHaM media were best suited for future metabolic flux mapping studies designed to probe the metabolic control points that determine resource partitioning between protein and oil. The SHaM results under conditions of excess C and N illustrate: 1) the protein content of the embryos increased from ~ 16 to 40% Dwt as the C: N mole ratio decreased; 2) the protein content of the embryos indicated a strong positive correlation with glutamine consumption rate, $R^2 = 0.68$; 3) starch content was linearly correlated to increased C: N ratios, $R^2 = 0.57$, as was the residual biomass fraction, estimated by mass balance, $R^2 = 0.77$; and 4) oil synthesis appeared to be saturated, as oil content was independent of C: N ratios. The data indicate that the partitioning of assimilates into protein was highly dependent on the supply of reduced organic nitrogen in the media. Furthermore, starch in SHaM embryos appears to represent a sink for excess carbon after the biosynthesis of other storage products, such as protein. The presented data show clear relationships between storage products and growth under carbon-limited regimes, as well as under conditions of excess carbon and nitrogen supply.

2 Materials and Methods

2.1 Plant material and culture conditions

Cultures of proliferative soybean somatic embryos *cv. Jack* were provided by DuPont Pioneer Hi-Bred and were maintained, by weekly subculture, in SB196 medium. SB196 medium is a modified form of MSD20 medium (Walker and Parrott, 2001), and contained FN Lite Halides, FN Lite P, B, Mo, Murashige-Skoog (MS) sulfate, MS Fe EDTA, B5 vitamins, KNO_3, L-asparagine, and 29 mM sucrose. In addition, the SB196 medium contained a moderately high concentration of synthetic auxin, 2, 4-dichlorophenoxyacetic acid (10 mg/l). Clusters of globular-stage embryos (approximately 10 mg) were used to initiate cultures by dropping them into 250-ml Erlenmeyer flasks containing 35 ml SB228 liquid media (Schmidt *et al.*, 2005). The SB228 medium contained FN Lite macro salts, MS micro salts, B5 vitamins, $CaCl_2$, L-methionine, 88 mM sucrose, 30 mM glutamine, and 165 mM sorbitol. Embryos were induced, maintained, and matured under a light intensity of 35 – 50 µmol photons m^{-2} s^{-1}, provided by cool-white fluorescent bulbs, and a 16-h photoperiod. Cultures were maintained at 26 °C and were continuously shaken at 130 rpm for two weeks to build biomass prior to experimentation.

After two weeks in culture, working in a laminar flow hood, clusters of SHaM embryos collected from 20 flasks derived from the same subculture cycle, were separated into batches of approximately 40 to 50 randomly chosen, uniformly sized embryos, with an initial fresh weight of 1.00 g (measured and recorded to a precision of 0.01g), unless stated otherwise. These were placed into 250-ml Erlenmeyer flasks containing 35 ml of SHaM media with the different sucrose and glutamine

concentrations to be tested (Table 1). The media in each flask was exchanged every three days, unless stated otherwise. Cultures were harvested after six days by pouring the contents of each culture into a sieve. The harvested embryos were rinsed with 100 ml distilled water, blotted dry, weighed (fresh weight determination) and frozen in liquid nitrogen prior to lyophilization at -50°C and 0.0158 mbar for 72 hours. The lyophilized embryos were weighed (dry weight (Dwt) determination) and were finely ground using a Geno/Grinder® (SPEX SamplePrep, Metuchen NJ, USA) prior to further analysis.

Initial studies testing the influence of media composition (ranges of sucrose concentration at fixed glutamine (30mM); ranges of glutamine concentration at fixed sucrose (88 mM); the influence of inorganic nitrogen (ammonium nitrate)) on SHaM embryo growth and composition were performed as described above except that: the cultures were initiated with 2.00 g (measured and recorded to a precision of 0.01g) of uniformly sized embryos. The cultures were harvested after 7 days and the media was not exchanged during the experimental culture period.

2.2 Relative Growth Rate

The growth rates of the SHaM embryos were normalized based on their relative linear growth rates. Growth was expressed as the dry mass added during a given growth interval, per unit of original dry mass at the beginning of the growth interval. The relative growth rate (day^{-1}) was calculated using the following equation:

$$\mu_{relative} = \frac{\left[\frac{X_{dayn} - X_{day0}}{t}\right]}{X_{day0}}$$

where, X_{dayn} is the tissue Dwtat final harvest, X_{day0} is the initial Dwt of the culture (see below), and t is the culture duration. To determine the initial Dwt of the culture, three 1.00 gram fresh weight batches of representative embryos were prepared at the same time that the experimental cultures were being initiated (see above). The embryos were placed into pre-weighed 50-ml centrifuge tubes and were frozen in liquid nitrogen prior to lyophilization, as described above. After lyophilization, the embryo Dwt was determined and used in the calculations described above.

2.3 Lipid and Protein Extraction

Approximately 100 mg of dry powdered embryo sample was extracted with 1 ml of n-hexane at 40°C for 1 hour; the extracts were centrifuged at high speed (13,200 rpm) for 10 minutes at room temperature. The process was repeated five times and after each extraction, the solvent containing the lipids were pooled into a pre-weighed glass tube (approximately 4.8 ml final volume) and dried for four days in a hood at room temperature. The mass of lipids after solvent evaporation were measured gravimetrically. The remaining hexane extracted biomass was dried and further extracted for protein in 600 µl of 200 mM phosphate buffer (pH = 7.2) containing 14 mM β-mercaptoethanol at 4°C for 20 minutes to dissolve and suspend the protein into solution. The extracts were centrifuged at high speed at 4°C for 15 minutes and the supernatant was transferred into a micro-centrifuge tube. The extraction was performed on the pellet two more times with 400 µl of the buffer (per extraction) and the supernatants were pooled with that from the initial extraction (approximately 1.2 ml final volume). Protein contents of the extracts were measured by Bradford Assay (Bradford, 1976). The soluble protein extraction also removed soluble sugars from

the samples. Aliquots of the protein extract were therefore mixed with 100% ethanol at a ratio of 1 ml of protein extract/1ml of 100% ethanol, vortex mixed and centrifuged for 2 minutes at high speed. The supernatants were run on an HPLC for soluble sugars analysis (see below).

2.4 Residual Soluble Sugar and Starch Extraction

Residual soluble sugars in the defatted/deproteinated- biomass pellets, were extracted into 1 ml of 80% aqueous ethanol in a water bath sonicator (Fisher Scientific FS110H Ultrasonic Cleaner) at 60°C for 20 minutes. The extraction was repeated four times and after each extraction, the solvent containing the soluble sugars were pooled into a pre-weighed glass tube (approximately 4.8 ml final volume) and dried for 3 days in a 40 °C oven, and the soluble sugar content in these fractions was measured gravimetrically. To obtain the total soluble sugar content for the tissue, the results of HPLC analysis for soluble sugars in the protein extracts were added to those obtained for the 80% ethanol extracts. The remaining pellets from each sample were placed in 1.5 ml of DDI water, covered with foil and autoclaved (liquid cycle at 121 °C and a pressure of 15 psi) for 30 minutes prior to starch digestion and extraction. Starch was digested in 1.5 ml of 100 mM citrate buffer (pH 5.0) containing amyloglucosidase at a ratio of 0.025 mg enzyme: 1 mg of tissue Dwt (approximately 2 µl of amyloglucosidase or 0.6 units of enzyme per sample). Samples were incubated overnight in a 30°C water bath. Starch content was quantified using a Glucose Assay Kit (Sigma, St. Louis, MO).

2.5 Residual Biomass Fraction Estimation

The percentage of the residual biomass fraction was estimated by subtraction of the measured sum of the biomass, i.e., protein (by combustion analysis) + oil + starch + soluble sugar + ash and an estimation of the DNA/RNA content (see below), from the original mass of tissue extracted. Ash content was determined by Thermo-Gravimetric Analysis (TGA) as described below. DNA/RNA content was assumed to be 5 percent on a Dwt basis (Stephanopoulos *et al.*, 1998).

2.6 Determination of Ash Content

Ash content was determined by complete combustion of the samples in a Thermo-gravimetric Analyzer (TGA 7) running the following temperature ramp program: ramp from room temperature to 100 °C and hold for 10 min; ramp to 200 °C at 10 °C/min; hold at 200 °C for 10 min; ramp to 400°C at 10°C/min; hold at 400 °C for 30 min; ramp to 700 °C at 10 °C/min; hold at 700 °C for 180 min; drop to, and hold at, room temperature until sample removal. The remaining residue in the crucible was taken as the ash content.

2.7 Sucrose, Glutamine Measurement, and Elemental Analysis

Sucrose and glutamine consumption rates from the media were determined by measuring the residual sucrose and glutamine in the culture media at the initiation of the experiments, in the media recovered at the three-day exchange, and at culture was harvested. Each media sample was filtered with a syringe filter (25mm, 0.45 µm, Nylon, XPERTEX®) prior to injection into an HPLC (Waters 1525 Binary Pump and 717^{plus} Auto Sampler fitted with an Aminex HPX-87K column, and a Waters 2414 RI detector). Samples were eluted with a water mobile phase at 50°C, at a flow rate of

0.35 ml/minute; each run was 80 minutes. By running the separation for 80 minutes, both sucrose and glutamine were detected in the same HPLC run. The HPLC run also detected sorbitol in the liquid medium. The concentration of soluble sugars from the protein extracts were analyzed as above, except the runs were 30 minutes long to detect sucrose, glucose, and fructose. A series of sucrose (15, 44, 88, 146, 176, and 234 mM), glutamine (5, 10, 20, 30, 40, and 60 mM), and sorbitol (50, 100, 150, 200 mM) standards were prepared and measured with the unknown samples. Standard curves were generated by plotting the area under the peaks versus their concentrations and were used to compute the concentrations of the unknown samples. In the protein extracts, glucose and fructose concentrations were very low and calibration curves were independently derived for these sugars.

Sucrose and glutamine consumption rates were calculated using the following equation:

$$\frac{dS}{dt} = \frac{\left(S_{day6} - S_{day0}\right) \times V}{\left(t_{day6} - t_{day0}\right) \times \left(X_{day6} - X_{day0}\right)}$$

where S is the media sucrose or glutamine concentration at the onset and end of the culture period, V is the volume of the media after three or six-days of culture, t is time, and X is the embryo Dwt.

The Dwt percentages of carbon, hydrogen, and nitrogen weremeasured by elemental analysis using the Perkin-Elmer Model 2400 Series II CHN&S Elemental analyzer (Chemical Instrument Facility, Chemistry Department at Iowa State

University). In the determinations of biomass used throughout this study, total protein content was calculated from the elemental analysis (% N x 6.25).

2.8 Determination of free amino acids

Approximately 20 mg of dried ground embryo material was weighed into a microcentrifuge tube (to a precision of 0.1mg) and 40 µl of norleucine (1mg/ml DI water) was added as an internal standard followed by methanol:chloroform:water (700 µl:300 µl:450 µl). The samples were vortexed for 1 min, centrifuged at 15,000 rpm for 10 min at 4 ^{o}C. A 300 µl aliquot of the upper polar phase was transferred into a new 1.5 ml microcentrifuge tube and dried in a vacuum concentrator (Vacufuge Plus, Eppendorf, Hauppauge NY, USA). After centrifugation at 15,000 rpm for 2 min, 30 µl of the solution was transferred into GC vials fitted with glass inserts. An equal volume (30 µl) of N-tert-butyldimethylsilyl-N-methyltrifluoroacetamide containing 1% tert-butyldimethylchlorosilane (Sigma-Aldrich, St. Louis) was added and the vials were capped tightly and incubated at 80 ^{o}C for 30 min in a heated block. The TBDMS-derivatized amino acids were analyzed by gas chromatograph on an Agilent 6890 fitted with an HP-5MS (30m x 0.25 mm i.d., 0.25 µm) column and a 5973 mass spectrometer detector. The temperature program was as follows: 70 ^{o}C for 1min, ramp at 15 ^{o}C/min to 130 ^{o}C followed by a ramp of 5 ^{o}C/min to 300 ^{o}C, which was held for 1min. Amino Acid Standards for thirteen amino acids (0.25 µmol/ml for each amino acid; Thermo Fisher Scientific) were diluted in series (0.25, 0.125, 0.063, 0.031, 0.016 µmol/ml) with DI water. The standard solutions were mixed with the internal standard solution, (1mg/ml, norleucine) at a 1:1 volume ratio, and were dried and derivatized, as described above, prior to GC-MS analysis. Based on the retention

times and response factors for mass fragments of the amino acids and the internal standard, intracellular free amino acids were identified, and quantified. In order to estimate the influence of the sample matrix (proteins and salts) on amino acid derivatization efficiency, three sets of samples were prepared with internal standards; 1) samples alone (extracted from cultures grown at high, medium and low C: N ratios); 2) the same samples spiked with the thirteen amino acid standards eat at a defined concentration of 0.1 µmol; 3) the thirteen amino acid standards at the same concentrations as in 2. Glutamate (Glu) and alanine (Ala) were the dominant amino acids present in the tissue extracts. The derivatization efficiency of Glu and Ala were 90.2 ± 8.5 % and 96.7 ± 8.5 %, respectively. These derivatization efficiency values were used to estimate the free amino acid content of the tissue.

2.9 Statistical Analysis

Data has been presented as the mean ± S.E. for each media condition treatment (n = 3). Initial sucrose and glutamine concentrations were converted to carbon-to-nitrogen (C: N) mole ratios for each of the media conditions (Table 1). Significant differences between the treatments were determined by ANOVA, using the multiple comparison method of Tukey-Kramer's Honestly Significant Difference (HSD). Least significant difference (LSD) values were calculated at the $\alpha = 0.05$ probability level (JMP v. 8.0.2, SAS Institute Inc. Cary NC, USA).

3 Results

3.1 Determining the Carbon-limited regime for SHaM embryos

Typical medium concentrations of the carbon and nitrogen sources for zygotic embryos cultured *in vitro* are designed to mimic the conditions experienced by

developing seed on the plant. The actual concentrations of sucrose and amino acids at the apoplastic interface between the maternal and filial tissues are extremely difficult to measure with any degree of certainly and reported values lie in the range of 3.4 to 200 mM sucrose and 10 to 44 mM amino acids (Gifford and Thorne, 1985; Hsu *et al.*, 1984; Pipolo *et al.*, 2004; Saravitz and Raper, 1995; Schmidt *et al.*, 2005). Therefore, in this study, we used a broad range of sucrose and glutamine concentrations to better define the influence of the carbon and nitrogen source concentrations on embryo growth and composition. The regime where the sucrose was not limiting growth was determined from experiments performed at a fixed glutamine concentration (30 mM) and sucrose concentrations ranging from 0 to 234 mM. The highest sucrose concentration in the medium tested was slightly high when compared to the measured sucrose concentrations reported for the soybean seed apoplast, 150 – 200 mM, (Gifford and Thorne, 1985) but was lower than the upper range of sucrose concentrations (73-292 mM) used by Thompson et al., (1977) for the *in vitro* culture of soybean zygotic embryos.As shown in Fig. 1A, the slope of the relative growth rate (day^{-1}) versus sucrose concentration at sucrose concentrations below 88 mM was steeper than at higher concentrations (Fig. 1A). This indicated that carbon was limiting below approximately 88 mM sucrose, an observation that was supported by a total absence of starch in the embryos cultured at sucrose concentrations below 88 mM(Fig. 1B). However, at the lowest media sucrose concentrations tested (0 and 44 mM), the soluble sugar content of the SHaM embryos was significantly greater than that in embryos cultured in media with higher sucrose concentrations (Fig. 1C). SHaM embryos with the highest soluble sugar concentrations (i.e., those cultured in

the 0-44 mM sucrose range) were darker green than those cultured at higher sucrose concentrations (data not shown). We speculate that under the limiting sucrose regime, the SHaM embryos may be capable of autotrophic synthesis of sugars via photosynthesis. It is apparent that photosynthetic synthesis of sugars did not result in starch accumulation in these tissues.

In a complementary study we investigated the influence of changing the glutamine concentration (from 0 – 37mM) at a fixed sucrose concentration (88 mM). Under these conditions relative growth rates were unaffected across the range of glutamine concentrations tested (Figure S1A). Glutamine concentrations also had little influence on the starch, oil, and soluble sugar contents of the embryos, which all remained relatively constant across the glutamine concentration range(Figures S1B and S1C). In contrast, protein contents were positively correlated to increasing glutamine concentrations in the media (Figure S1B) and the residual biomass fraction showed a negative correlation (Figure S1C).

3.2 Linear Relative Growth Rate of SHaM embryos

Having established a sucrose concentration below which culture growth was significantly reduced, the influence of a range of carbon (sucrose) to nitrogen (glutamine) [C: N] mole ratios on the partitioning of resources between protein and oil was investigated. The protocol was also slightly modified in that cultures were initiated with only 1 g of embryonic tissue (c.f., 2 g in the initial studies) so as to reduce the degree of resource depletion during the culture period; the culture solutions were also changed after 3 days and cultures were harvested after 6 days). Sucrose concentrations from 88 to 234 mM, in combination with glutamine

concentrations of 20 to 60 mM, were selected for further experimentation. Table 1 shows the ratio of moles of C, provided by sucrose, to the moles of N, provided by glutamine in the initial SHaM media.

The linear relative growth rates of SHaM embryos for 20 different C: N mole ratios are shown in Fig. 2. The average relative growth rate under all media conditions for SHaM embryos was 0.61 ± 0.07 (day^{-1}) during the 6-day culture period. The lowest relative growth rate, 0.44 ± 0.01 (day^{-1}), was at a C: N mole ratio in the media of 26.40. This low relative growth rate was at a media condition of 88 mM sucrose and 20 mM glutamine, which is consistent with the fact that these conditions are at the boundary of the carbon and nitrogen-limited regime. The relative growth rate data indicate that for the most part, the SHaM embryos were not under carbon or nitrogen limitation.

3.3 Effect of Sucrose and Glutamine Supplements on Biomass Composition

The protein content of the embryos increased from ~ 16 to 40 % Dwt as the C: N mole ratio decreased from 70.20 to 8.80, with the regression coefficient $R^2 = 0.79$ (Figure S2A). This result is consistent with the role of glutamine as a main reduced nitrogen source available for protein biosynthesis. The base media used for these studies contained 28 mM inorganic nitrogen, in the form of potassium nitrate. To test whether the inorganic nitrogen, could influence SHaM growth and composition, an experiment was performed in which the glutamine was replaced with ammonium nitrate at 5, 15, 30, and 40 mM; sucrose concentration was fixed at 88 mM. The relative growth rate of the embryos was not influenced by the inorganic nitrogen concentration (Figure S3A). However, in contrast to the results where nitrogen was

supplied in the form of glutamine, inorganic nitrogen concentrations had little influence on embryo composition (Figure S3B and C). Notably, protein contents were unaffected.

Bradford assays were performed on a selection of the samples, as an independent estimate of protein content. The results, along with the protein content estimates from elemental analysis of the same tissues, are given in Table 2. The Bradford assay (which only measures soluble proteins) were only 40-50% of those determined by elemental analysis but the trend, i.e., increases in protein content in response to decreasing C: N ratios, were equally apparent. One concern when using combustion analysis for determining protein content is that the method reports total nitrogen content for a sample and, as such, includes both total protein and any other nitrogen containing molecules, both organic and inorganic. Because our treatments involved supplying some of the cultures with very high levels of glutamine in the media, it was important to determine the contribution of endogenous free amino acid pools to our overall measurement of tissue nitrogen content. Analysis of the free amino acid pools show that the embryos contained significant levels of both glutamate (Glu) and alanine (Ala).All other amino acids detected were at concentrations that were two orders of magnitude lower (data not shown). Estimates of the contributions of the Glu and Ala pools to the total tissue percentage N measurements are given in Table 3. The data show that total protein contents were overestimated by approximately 20 ± 5 %.

The influence of sucrose and glutamine concentrations in the media on the oil content of SHaM embryos is shown in Fig. 3A. The oil content across the C: N ratios tested in these experiments were 5 to 8 %, on a Dwt basis. The oil content of SHaM

embryos therefore appears to have no correlation to the media C: N mole ratios, and is therefore in stark contrast to protein biosynthesis. To examine whether the oil and protein content of SHaM embryos were inversely related across the C: N ratios tested, the oil and protein contents were plotted against each other (Figure S4). Although the slope from a linear fit was -1.82, which indicates that a 1% reduction in oil content would lead to a 1.8% increase in protein content, similar to the rule of thumb reported in Clemente and Cahoon (2009), the correlation coefficient was $R^2 = 0.13$ (Figure S4) and is therefore insignificant in our system.

The starch contents in SHaM embryos in response to media composition are shown in Fig. 3A. The data indicated that there was a linear correlation between starch content and C: N mole ratios, with the linear correlation coefficient $R^2 = 0.57$ (Figure S2B). Moreover, the level of starch in the SHaM embryos decreased as the glutamine concentration increased at the four higher sucrose concentrations tested, (117 to 234 mM; data not shown). As shown above, on a Dwt percentage basis, protein content decreased with increasing media C: N mole ratios (Fig. 3A). Therefore, the starch and protein content of SHaM embryos were inversely related with a linear correlation coefficient, $R^2 = 0.59$ (data not shown).

The response of endogenous soluble sugar pools to media sucrose and glutamine concentrations is shown in Fig. 3B. The total soluble sugar content (from the ethanol and protein extractions combined) ranged between 14 and 20 % on a Dwt basis. There was no discernible correlation between the soluble sugars in the SHaM embryos and the media C: N mole ratios (Fig. 3B).

3.4 Closing the mass balance: residual biomass of SHaM embryos

To complete the biomass composition balance, the residual biomass fraction was estimated based on mass balance (see Materials and Methods section). The percentage residual biomass fraction was calculated by subtracting the sum of the biomass composition [protein (by elemental analysis) + oil + starch + soluble sugar + ash + DNA/RNA] from the mass of tissue extracted. The residual biomass fraction increased as the C: N mole ratio increased with the linear correlation coefficient R^2 = 0.77, as shown in Figure S2C. Protein and residual biomass of SHaM embryos were inversely related with a linear correlation coefficient, R^2 = 0.91 (Figure S2D). Preliminary studies showed that our lipid, and soluble sugar extractions were exhaustive. We are therefore confident that little residual oil, soluble sugar or protein (measured by combustion analysis) lead to errors in the calculation of the residual biomass fraction.

3.5 Glutamine and Sucrose uptake rate of SHaM embryos

The 3D graph of the glutamine consumption rate provides a clear indication that the SHaM embryos consumed more glutamine, as media glutamine concentrations increased from 20 to 60 mM, and the concentration of sucrose in the media decreased from 234 mM to 88 mM (Figure S5). Measurements of endogenous free amino acids pools show clearly that the majority of the glutamine taken up by the embryos was converted to protein. The protein content of the SHaM embryos indicated a strong positive correlation with the glutamine consumption rate, R^2 = 0.68, Figure S6.

The sucrose consumption rate was more complex. No significant correlation between sucrose consumption rate and the media C: N mole ratio was observed (data not shown). In our studies, the medium was replaced every three days to ensure that

the nutrient supplies were constant for carbon and nitrogen sources throughout the experimental period. The percentages of sucrose depletion under the 20 media conditions were examined for day 3 and day 6 cultures. The data indicated that after 3 and 6 days of culture, the percentages of sucrose depletion were 19.2 ± 8.5 and 29.0 ± 9.7, respectively (data not shown).

4 Discussion

In this study SHaM embryos were cultured in media with varying carbon-to-nitrogen mole ratios. To our knowledge, this is the first reported study of resource partitioning between storage products under various C: N ratios for soybean somatic embryos. We hypothesized that at constant growth rates, sucrose and glutamine in excess of that needed to support growth and maintenance metabolism would be directed to the synthesis of the storage products, protein, oil, and carbohydrates and that the proportional distribution would be influenced by the C: N mole ratios of the media. The information from this study also allowed the selection of a set of media conditions to be used for future metabolic flux mapping studies that are designed to probe metabolic control points regulating the partitioning of storage reserves. Such data will allow direct comparisons to be made to the fluxes previously reported for soy zygotic embryos (Allen *et al.*, 2009; Iyer *et al.*, 2008; Sriram *et al.*, 2004), including soy zygotic embryos grown under different C: N ratios (Allen and Young, 2013), in order to assess the relevance of measurements made in this model system to processes occurring during seed development in-planta.

For sucrose concentrations below 88 mM, and a constant glutamine concentration of 30 mM, (the recommended concentrations of these two components in standard

SHaM media (Schmidt *et al.*, 2005)) the SHaM embryos were carbon-limited. The relative growth rates decreased sharply in media with sucrose levels below 88 mM, Fig. 1A, but did not appear to be influenced by increases in the sucrose concentration above that amount, a result that is consistent with the observations of Egli and Bruening (2001) and Thompson et al., (1997) with soybean zygotic embryos. Interestingly, Thorn and Gifford (1985) measured the sucrose concentration in the tissue water of developing soybean cotyledons to fall in the 83-93 mM range during a diurnal cycle, a value that was significantly reduced (to 23 mM) only after defoliation of the mother plant the morning before the cotyledons were harvested. At media sucrose concentrations below 88 mM, the SHaM embryos did not accumulate starch. From these results and previous studies on soybean embryos cultured *in vitro* (Gifford and Thorne, 1985; Hsu *et al.*, 1984; Pipolo *et al.*, 2004; Saravitz and Raper, 1995; Schmidt *et al.*, 2005), an experimental scheme for testing the influence of C: N mole ratios on resource partitioning in SHaM embryos was established. As postulated, under excess carbon and nitrogen in the liquid media (concentrations equal or greater than 88 mM sucrose and 30 mM glutamine) the relative growth rates of the SHaM embryos were constant across the range of C: N mole ratios tested (8.80 to 70.20) during the six-day culture period, indicating the presence of an optimal supply of carbon and nitrogen sources for growth (Fig. 2).

For a meaningful comparison of the growth rate results between somatic embryos and zygotic embryos, all growth data were expressed as relative growth rates (Fig. 2). Under optimal conditions the average relative growth rate of SHaM embryos across the broad range of C: N mole ratios tested here was 0.61 ± 0.07 (day^{-1}). This

compared well with a value of 0.52 ± 0.10 (day^{-1}) measured for zygotic embryos of the same cultivar cultured *in vitro* at 146 mM sucrose and 37 mM glutamine (unpublished data). Further, the SHaM embryo relative growth rates were comparable to those of two different cultivars of zygotic embryos: *cv. Evans* (cultured *in vitro* at 146 mM sucrose and 37 mM glutamine) with a value of 0.45 ± 0.07 (day^{-1})(Iyer *et al.*, 2008), and *cv. Jack* (cultured *in vitro* at C: N ratios from 6 to 91) with an average value of 0.67 ± 0.10 (day^{-1})(Allen and Young, 2013).

In addition to their roles in supporting growth and as the principal precursors for storage product synthesis (Thompson *et al.*, 1977; Haga and Sodek, 1987) sucrose and glutamine have also been shown to be signaling molecules triggering storage-associated processes in plant sink tissues (Haga and Sodek, 1987; Koch, 2004; Koch, 1996; Smeekens, 2000; Thompson *et al.*, 1977). Our data clearly show that the partitioning of assimilates into protein was strongly influenced by the C: N mole ratio in the media, a relationship that had an $R^2 = 0.79$ (Fig. 3A; Figure S2A) and was therefore highly dependent on the glutamine concentration in the media. Similar results have been reported in, *in vitro* studies of soybean zygotic embryos, where increasing glutamine concentrations in the media resulted increased protein contents (Allen and Young, 2013; Pipolo *et al.*, 2004; Saravitz and Raper, 1995). In our study,the total protein content also correlated strongly with glutamine consumption rate, with an $R^2 = 0.68$ (Figure S6). When inorganic nitrogen (ammonium nitrate) was used to replace glutamine, embryo growth was little influenced across the concentration range tested but higher levels of inorganic nitrogen did not lead to increased protein contents (Figure S3). Taken together, these results indicate that

protein biosynthesis in SHaM embryos is extremely dependent on the supply of reduced organic nitrogen in the media. The conversion of glutamine to the other amino acids required for storage protein synthesis demands carbon skeletons derived from the metabolism of sucrose imported from the phloem (Miflin and Lea, 1977). In the case of glutamine, its amide-N-group must be donated to a suitable C-skeleton acceptor molecule, such as 2-oxoglutarate (a product of TCA cycle in the mitochondria), which can then be used in the formation of other amino acids (Weigelt *et al.*, 2008). Our data clearly show that as glutamine levels are increased (C: N mole ratios are decreased), protein synthesis draws carbon away from the non-structural (starch; Figure S2B) and structural (residual biomass; Figure S2C) fractions but does not influence the pool sizes of soluble sugars (data not shown). The negative correlation between the protein content and the structural carbohydrate pools is of significant interest as it is consistent with the observation of Somers and his colleagues who have demonstrated a negative relationship between the oil plus protein content and the fiber fraction in mature soybeans (Stombaugh *et al.*, 2000; Stombaugh *et al.*, 2003).

In a previous *in vitro* physiological study, Pipolo et al (2004), using zygotic embryos from a different soybean cultivar (*cv. Williams 82*) have reported an inverse relationship between oil and protein contents in response to different glutamine concentrations (20, 40, 60 and 80 mM) in the media, at a fixed sucrose concentration of 204 mM. The slope of the protein to oil content regression graph was -0.19, which, as the authors commented was significantly less than the expected value of -0.82, predicted from theoretical calculations of the conversion of photosynthate to protein

and oil mass. The authors conclude that the negative relationship between protein and oil in soybean seed is not simply a tradeoff between the energy requirements for protein and oil synthesis, a conclusion that is consistent with our findings where no clear relationship between protein and oil accumulation was apparent. In our studies when the sucrose concentration in the media was above 88mM, oil levels remained in the range from 5 – 8 % Dwt, and showed no correlation to C: N mole ratios (Fig. 3A). The oil levels observed in our experiments are in good agreement with the oil content of somatic embryos reported by other researchers (He *et al.*, 2011; Li *et al.*, 2010; Schmidt *et al.*, 2005). Although a linear regression line, derived from a plot of our oil and protein data, showing a 1% reduction in oil content lead to an apparent 1.82% increase in protein content, a value that was close to the 2% rule of thumb suggested for soybean by Clemente and Cahoon (2009), the correlation was very weak (Figure S4). Therefore, carbon that was in excess of that needed to synthesize protein, and for maintenance metabolism, did not go to oil synthesis but instead appeared to be directed to the starch and residual biomass pools. Therefore, in this model system the capacity to synthesize oil appears to be fixed and is not influenced by the C: N mole ratios supplied. Examination of the data reported for soy zygotic embryos (Allen and Young, 2013)reveals that the oil content was static for the broad range of C: N ratios tested. This observation is in good agreement with our results.

Our data provides evidence for the physiological role of starch accumulation in SHaM embryos. As observed in the carbon-limited regime, under media conditions containing less than 88 mM sucrose, there was no starch detected in the embryos at harvest (Fig. 1B). At sucrose concentrations above 88mM, starch did accumulate, but

protein synthesis appeared to be a more competitive sink for the available carbon pool. Therefore, starch appears to represent a sink for excess carbon after the biosynthesis of other storage products, such as protein. However, our results also indicate that starch does not serve as a temporary carbon storage reserve that can be readily used for the biosynthesis of oil, and is therefore in contrast to what has been reported for developing rape seed (Norton and Harris, 1975).

Soluble sugars such as sucrose, glucose, fructose, myo-inositol, raffinose, and stachyose are accumulated in developing soybean seeds (Wilson, 1995) and somatic embryos (Chanprame *et al.*, 1998). In our studies, sucrose in the media in excess of that needed for growth had no dramatic effect on the soluble sugar content (Fig. 3C). The data agree with field studies of zygotic embryos (Hymowitz *et al.*, 1972; Wilcox and Shibles, 2001), where soluble sugars had a weak interaction or were not associated with seed yield or protein, and oil content. Nonetheless, under carbon-limited conditions (0 to 44 mM, at constant glutamine concentration; 30 mM), the low sucrose concentrations did result in accumulation of soluble sugars in the tissues with levels rising to more than 25 %, on a Dwt basis (Fig. 1C). In addition to the increased soluble sugar content, the color of the embryos was observed to change from light to darker green (data not shown). One possible explanation for this is that under severe carbon-limitation, embryos synthesized their own sugars via photosynthesis. Photosynthesis is active in early phases of soybean seed development (Borisjuk *et al.*, 2005). Our data indicate that the soluble sugar pools are highly regulated in embryonic tissues.

The weak correlation between oil and protein content in embryos suggested that oil and protein biosynthesis were independent at sucrose concentrations in excess of 88mM. These results contrast with our observations under sucrose limitation (at 30 mM glutamine), as both protein and oil were inversely related and there was a complete absence of starch (Fig. 1B). These results may be explained by the role that sucrose plays developing seeds. Sucrose is the primary translocated carbohydrate, and it plays an essential role in carbon partition for the biosynthesis of protein and oil (Smith *et al.*, 1989). As mentioned above, the starch level is strongly dependent on the sucrose concentration in the media. More importantly, the SHaM embryos used in this investigation clearly demonstrated that media compositions (in a range that does not adversely influence growth) had a relatively neutral influence on oil content, whereas protein and starch content were strongly influenced by media C: N mole ratios. Indeed, cytosolic amino acid and protein biosynthesis were shown to be independent from plastid fatty acid synthesis in *B. napus* in mid-storage phase despite the potential exchange of shared metabolic intermediates between the two pathways (Schwender and Ohlrogge, 2002).

Table S1 shows comparisons between published study focused on the influence of C: N ratios on assimilate partition in soybean zygotic embryos (Allen and Young, 2013) and our SHaM data. The data clearly show that SHaM embryos respond to changes in C: N ratios in similar manner to zygotic embryos and provide further support of their relevance as a model system for developing soybean seed. Furthermore, the experimental data reported in this study was highly reproducible, indicating the metabolic uniformity and growth characteristics of soybean somatic

embryos cultured in SHaM media. Therefore, SHaM embryos, which based on their composition and transcript profiling data (not shown) represent zygotic embryos in the early to mid-storage phase of development, are attractive as a means to test genotype-phenotype relationships in transgenic research programs. They are therefore an ideal test bed for the application of metabolic flux analysis studies designed to test hypotheses regarding carbon partitioning.

Acknowledgements

The authors would like to thank Christina M. Kostow for providing proliferative soybean somatic embryos and for providing the techniques for maintaining them in culture. Additionally the authors thank the editors and anonymous reviewers for their helpful comments and suggestions that have improved this manuscript. This work was supported by DuPont Pioneer Hi-Bred.

References

Allen DK, Ohlrogge JB, Shachar-Hill Y. 2009. The role of light in soybean seed filling metabolism. *Plant Journal* 58, 220-234.

Allen DK, Young JD. 2013. Carbon and Nitrogen Provisions Alter the Metabolic Flux in Developing Soybean Embryos. *Plant Physiology* 161, 1458-1475.

Borisjuk L, Nguyen TH, Neuberger T, et al. 2005. Gradients of lipid storage, photosynthesis and plastid differentiation in developing soybean seeds. *New Phytologist* 167, 761-776.

Bradford MM. 1976. Rapid and Sensitive Method for Quantitation of Microgram Quantities of Protein Utilizing Principle of Protein-Dye Binding. *Analytical Biochemistry* 72, 248-254.

Chanprame S, Kuo TM, Widholm JM. 1998. Soluble Carbohydrate Content of Soybean [Glycine max (L.) Merr.] Somatic and Zygotic Embryos during Development. *In Vitro Cellular & Developmental Biology. Plant* 34, 64-68.

Clemente TE, Cahoon EB. 2009. Soybean Oil: Genetic Approaches for Modification of Functionality and Total Content. *Plant Physiology* 151, 1030-1040.

Finer JJ, McMullen MD. 1991. Transformation of Soybean Via Particle Bombardment of Embryogenic Suspension Culture Tissue. *In Vitro Cellular and Developmental Biology Plant* 27P, 175-182.

Gifford RM, Thorne JH. 1985. Sucrose Concentration at the Apoplastic Interface Between Seed Coat and Cotyledons of Developing Soybean Seeds. *Plant Physiology* 77, 863-868.

Haga KI, Sodek L. 1987. Utilization of Nitrogen-Sources by Immature Soybean Cotyledons in Culture. *Annals of Botany* 59, 597-601.

Hartwig EE, Kilen TC. 1991. Yield and Composition of Soybean Seed from Parents with Different Protein, Similar Yield. *Crop Science* 31, 290-292.

Hayati R, Egli DB, CraftsBrandner SJ. 1996. Independence of nitrogen supply and seed growth in soybean: Studies using an in vitro culture system. *Journal of Experimental Botany* 47, 33-40.

He Y, Young T, Clark K, Kleppinger-Sparace K, Bridges W, Sparace S. 2011. Developmental profile of storage reserve accumulation in soybean somatic embryos. *In Vitro Cellular & Developmental Biology - Plant* 47, 725-733.

Hsu FC, Bennett AB, Spanswick RM. 1984. Concentrations of Sucrose and Nitrogenous Compounds in the Apoplast of Developing Soybean Seed Coats and Embryos *Plant Physiology* 75, 181-186.

Hymowitz T, Collins FI, Walker WM, Panczner J. 1972. Relationship between content of oil, protein, and sugar in soybean seed. *Agronomy Journal* 64, 613-616.

Iyer VV, Sriram G, Fulton DB, Zhou R, Westgate ME, Shanks JV. 2008. Metabolic flux maps comparing the effect of temperature on protein and oil biosynthesis in developing soybean cotyledons. *Plant Cell and Environment* 31, 506-517.

Kinney AJ. 1996. Development of genetically engineered soybean oils for food applications. *Journal of Food Lipids* 3, 273-292.

Koch K. 2004. Sucrose metabolism: regulatory mechanisms and pivotal roles in sugar sensing and plant development. *Current Opinion in Plant Biology* 7, 235-246.

Koch KE. 1996. Carbohydrate-modulated gene expression in plants. *Annual Review of Plant Physiology and Plant Molecular Biology* 47, 509-540.

Li ZS, Moon BP, Xing AQ, Liu ZB, McCardell RP, Damude HG, Falco SC. 2010. Stacking Multiple Transgenes at a Selected Genomic Site via Repeated Recombinase-Mediated DNA Cassette Exchanges. *Plant Physiology* 154, 622-631.

Lu C, Napier JA, Clemente TE, Cahoon EB. 2011. New frontiers in oilseed biotechnology: meeting the global demand for vegetable oils for food, feed, biofuel, and industrial applications. *Current Opinion in Biotechnology* 22, 252-259.

Meyer K, Hitz WD, Yadav NS, Damude HG. 2012. DGAT genes from Yarrowia lipolytica for increased seed storage lipid production and altered fatty acid profiles in soybean. E I du Pont de Nemours and Company. US Patent No US 8143473.

Miflin BJ, Lea PJ. 1977. Amino Acid Metabolism. *Annual Review of Plant Physiology* 28, 299-329.

Nishizawa K, Ishimoto M. 2009. Maturation of somatic embryos as a model for soybean seed development. *Plant Biotechnology* 26, 543-550.

Nishizawa K, Takagi K, Teraishi M, Kita A, Ishimoto M. 2010. Application of somatic embryos to rapid and reliable analysis of soybean seed components by RNA interference-mediated gene silencing. *Plant Biotechnology* 27, 409-420.

Norton G, Harris JF. 1975. Compositional Changes in Developing Rape Seed (Brassica-Napus L) *Planta* 123, 163-174.

Pipolo AE, Sinclair TR, Camara GMS. 2004. Protein and oil concentration of soybean seed cultured in vitro using nutrient solutions of differing glutamine concentration. *Annals of Applied Biology* 144, 223-227.

Saravitz CH, Raper CD. 1995. Responses to Sucrose and Glutamine by Soybean Embryos Growth *In-Vitro*. *Physiologia Plantarum* 93, 799-805.

Schmidt MA, Tucker DM, Cahoon EB, Parrott WA. 2005. Towards normalization of soybean somatic embryo maturation. *Plant Cell Reports* 24, 383-391.

Schmutz J, Cannon SB, Schlueter J, et al. 2010. Genome sequence of the palaeopolyploid soybean. *Nature* 463, 178-183.

Schwender J, Ohlrogge JB. 2002. Probing in vivo metabolism by stable isotope labeling of storage lipids and proteins in developing Brassica napus embryos. *Plant Physiology* 130, 347-361.

Schwender J, Ohlrogge JB, Shachar-Hill Y. 2003. A flux model of glycolysis and the oxidative pentosephosphate pathway in developing Brassica napus embryos. *Journal of Biological Chemistry* 278, 29442-29453.

Sinclair TR, de Wit CT. 1975. Photosynthate and Nitrogen Requirements for Seed Production by Various Crops. *Science* 189, 565-567.

Smeekens S. 2000. Sugar-induced signal transduction in plants. *Annual Review of Plant Physiology and Plant Molecular Biology* 51, 49-81.

Smith AJ, Rinne RW, Seif RD. 1989. Phosphoenolpyruvate Carboxylase and Pyruvate Kinase Involvement in Protein and Oil Biosynthesis during Soybean Seed Development. *Crop Sci.* 29, 349-353.

Sriram G, Fulton DB, Iyer VV, Peterson JM, Zhou RL, Westgate ME, Spalding MH, Shanks JV. 2004. Quantification of compartmented metabolic fluxes in developing soybean embryos by employing Biosynthetic ally directed fractional C-13 labeling, C-13, H-1 two-dimensional nuclear magnetic resonance, and comprehensive isotopomer balancing. *Plant Physiology* 136, 3043-3057.

Stephanopoulos GN, Aristidou AA, Nielsen J. 1998. Metabolic Flux Analysis. *Metabolic Engineering*. San Diego: Academic Press, 309-351.

Stombaugh SK, Jung HG, Orf JH, Somers DA. 2000. Genotypic and environmental variation in soybean seed cell wall polysaccharides. *Crop Science* 40, 408-412.

Stombaugh SK, Orf JH, Jung HG, Somers DA. 2003. Relationships between soybean seed cell wall polysaccharides, yield, and seed traits. *Crop Science* 43, 571-576.

Thompson JF, Madison JT, Muenster AME. 1977. *In Vitro* Culture of Immature Cotyledons of Soybean (*Glycine-Max-L-Merr*) *Annals of Botany* 41, 29-39.

Walker DR, Parrott WA. 2001. Effect of polyethylene glycol and sugar alcohols on soybean somatic embryo germination and conversion. *Plant Cell Tissue and Organ Culture* 64, 55-62.

Weigelt K, Kuster H, Radchuk R, Muller M, Weichert H, Fait A, Fernie AR, Saalbach I, Weber H. 2008. Increasing amino acid supply in pea embryos reveals specific interactions of N and C metabolism, and highlights the importance of mitochondrial metabolism. *Plant Journal* 55, 909-926.

Wilcox JR, Shibles RM. 2001. Interrelationships among seed quality attributes in soybean. *Crop Science* 41, 11-14.

Wilson LA. 1995. Soy foods. In D. R. Erickson (ed.) Practial handbook of soybean processing and utilization. AOCS Press, Champaign, IL and United Soybean Board, St. Louis, MO., 428-459.

Figures and Legends

Figure 1: Relative growth rate (day^{-1}) (A); and compositional analysis of: Protein, Oil, and Starch (B); Soluble sugars and Residual biomass content (C) in soybean somatic embryos *cv. Jack* as a function of the initial sucrose concentrations in the SHaM media; note the media was not changed during the 7 day culture period. Initial

sucrose concentrations ranged from 0 to 234 mM. The initial concentration of glutamine was 30 mM. Below 88 mM sucrose (indicated by dashed line), SHaM embryos were carbon limited, and starch was not detected in these cultures at harvest.

Figure 2: The relative growth rates of soybean somatic embryos *cv. Jack*, day^{-1}, with different initial Carbon-to-Nitrogen (C: N) mole ratios. The embryos were cultured in SHaM media for 6 days with a media change on day 3. Error bars represent standard error (n = 3). The diamonds show the confidence limits ($\alpha = 0.05$) for each mean (determined by analysis of Variance (ANOVA) of the three biological replicates at each media condition).Tukey-Kramer's Honestly Significant Difference analysis show that the relative growth rates, within each box, did not differ significantly at the $\alpha = 0.05$ level.

Figure 3: Protein, oil, starch (A); soluble sugars and residual biomass (B), contents, all expressed on a dry weight basis, for soybean somatic embryos *cv. Jack* cultured in SHaM media with different initial Carbon-to-Nitrogen (C: N) mole ratios for 6 days. The media was replaced after 3 days of culture.

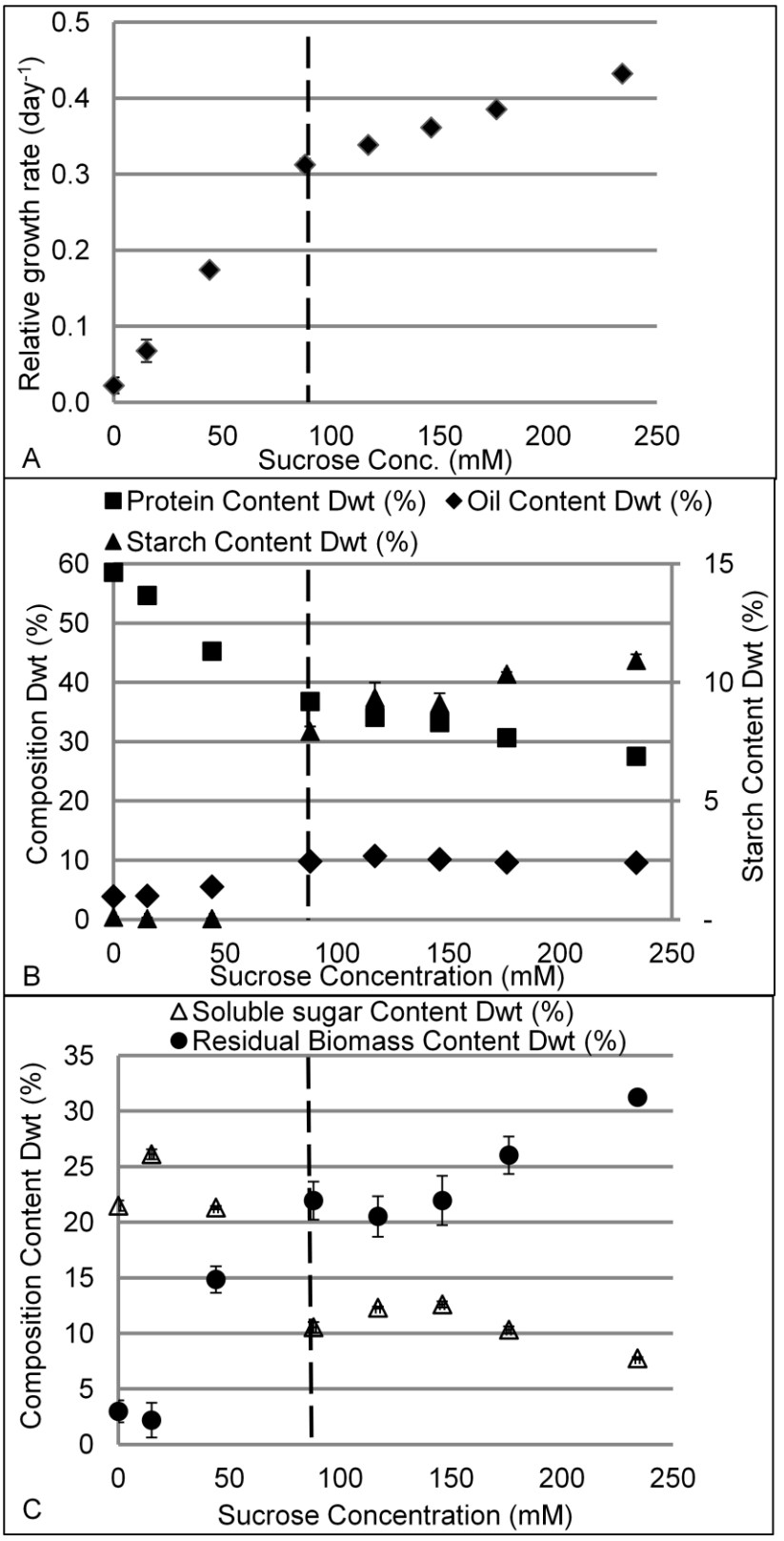

Relative growth rate (day-1)
0.5
0.4
0.3
0.2
0.1
0.0
0
50
100
150
200
250
Sucrose Conc. (mM)
A
Protein Content Dwt (%)
Oil Content Dwt (%)
Starch Content Dwt (%)
Composition Dwt (%)
60
50
40
30
20
10
0
15
10
5
-
Starch Content Dwt (%)
Sucrose Concentration (mM)
B
Soluble sugar Content Dwt (%)
Residual Biomass Content Dwt (%)
Composition Content Dwt (%)
35
30
25
20
15
10
5
0
Sucrose Concentration (mM)
C

Figure 1

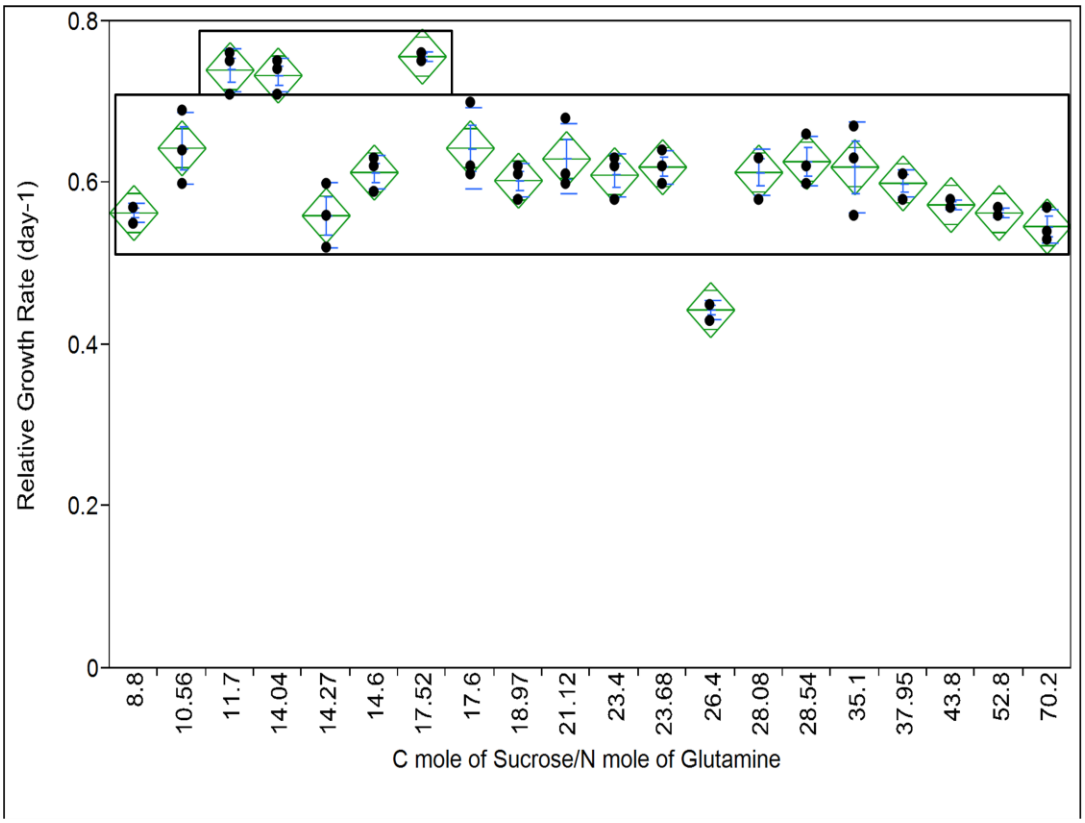

Relative Growth Rate (day-1)
0.8
0.6
0.4
0.2
0
8.8
10.56
11.7
14.04
14.27
14.6
17.52
17.6
18.97
21.12
23.4
23.68
26.4
28.08
28.54
35.1
37.95
43.8
52.8
70.2
C mole of Sucrose/N mole of Glutamine

Figure 2

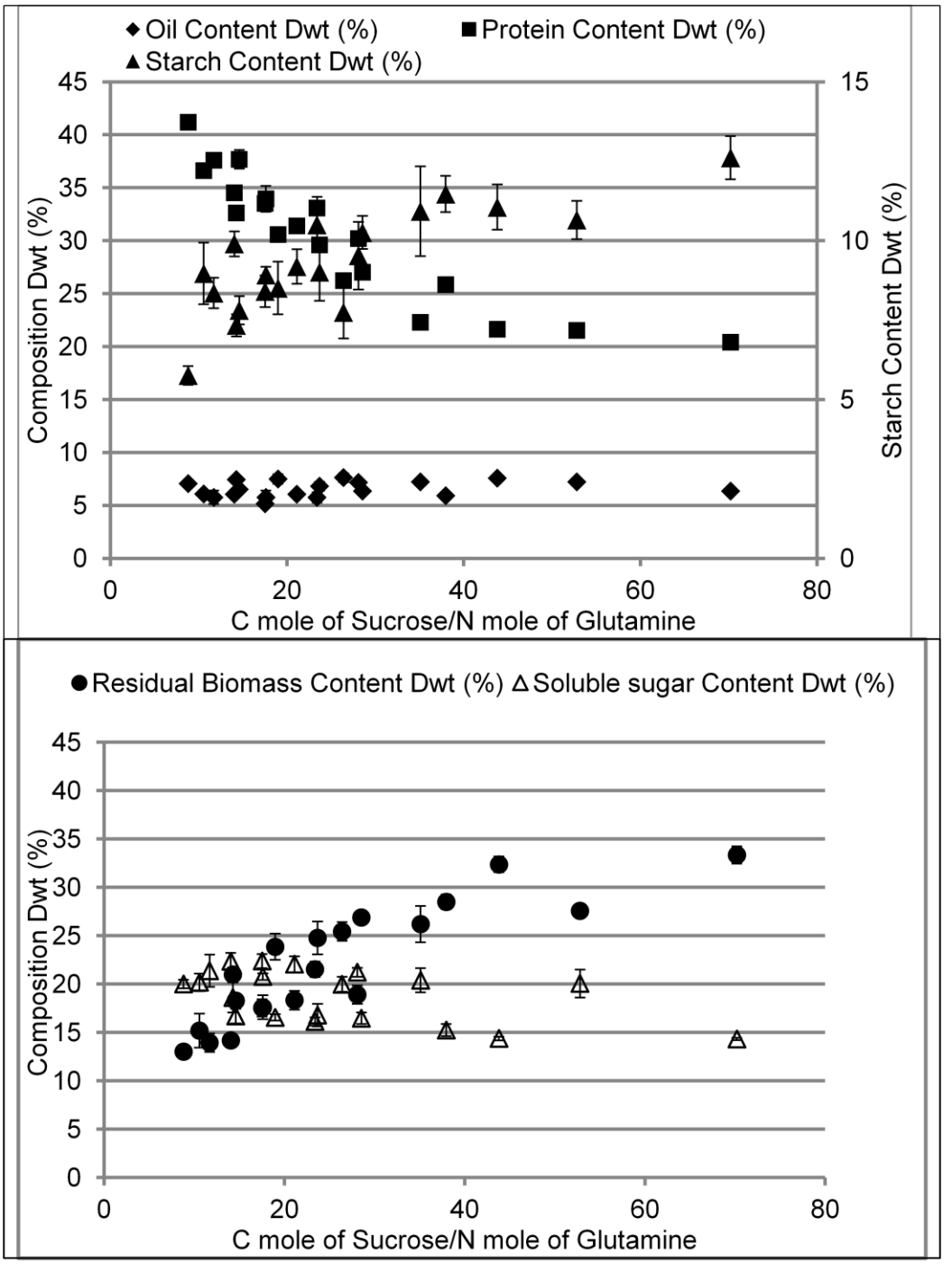

Oil Content Dwt (%)
Protein Content Dwt (%)
Starch Content Dwt (%)
Composition Dwt (%)
Starch Content Dwt (%)
C mole of Sucrose/N mole of Glutamine
Residual Biomass Content Dwt (%)
Soluble sugar Content Dwt (%)
Composition Dwt (%)
C mole of Sucrose/N mole of Glutamine

Figure 3

Table 1: Media sucrose and glutamine concentrations and the mole ratios of Carbon-Nitrogen used in this study.

C moles of Sucrose/N moles of Glutamine (C: N mole ratio)				
	Glutamine Conc. (mM) →			
Sucrose Conc. (mM) ↓	***20***	***37***	***50***	***60***
88	26.40	14.27	10.56	8.80
117	35.10	18.97	14.04	11.70
146	43.80	23.68	17.52	14.60
176	52.80	28.54	21.12	17.60
234	70.20	37.95	28.08	23.40

Table 2: Measurement of protein content, via elemental analysis and Bradford Assay, of soybean somatic embryos *cv. Jack* cultured under a selection of C: N mole ratios.

C:N mole ratio	Elemental Analysis	Bradford Assay
8.80	41.2 ± 0.5	16.4 ± 1.1
23.68	29.6 ± 0.7	13.4 ± 0.8
52.80	21.5 ± 0.3	11.4 ± 0.4

Table 3: Estimates of the apparent total nitrogen content, of soybean somatic embryos *cv. Jack* cultured under a selection of C: N mole ratios, contributed by endogenous free amino acids. Estimates of their possible contributions to the total protein contents measured by combustion analysis are given.

C:N mole ratio	% N from Alanine (Ala)	% N from Glutamate (Gln)	Apparent contribution to protein content [% N (Ala + Gln) * 6.25]	Total protein via Elemental Analysis
8.80	0.41 ± 0.11	1.33 ± 0.04	10.9 ± 0.6	41.2 ± 0.5
14.04	0.47 ± 0.01	0.81 ± 0.08	8.0 ± 0.4	34.5 ± 0.4
23.68	0.27 ± 0.01	0.80 ± 0.06	6.7 ± 0.4	29.6 ± 0.7
28.54	0.25 ± 0.02	0.73 ± 0.05	6.1 ± 0.3	27.0 ± 0.3
35.10	0.16 ± 0.01	0.32 ± 0.04	3.0 ± 0.3	22.3 ± 0.2
70.20	0.10 ± 0.01	0.32 ± 0.02	2.6 ± 0.2	20.4 ± 0.1

Supplementary Data

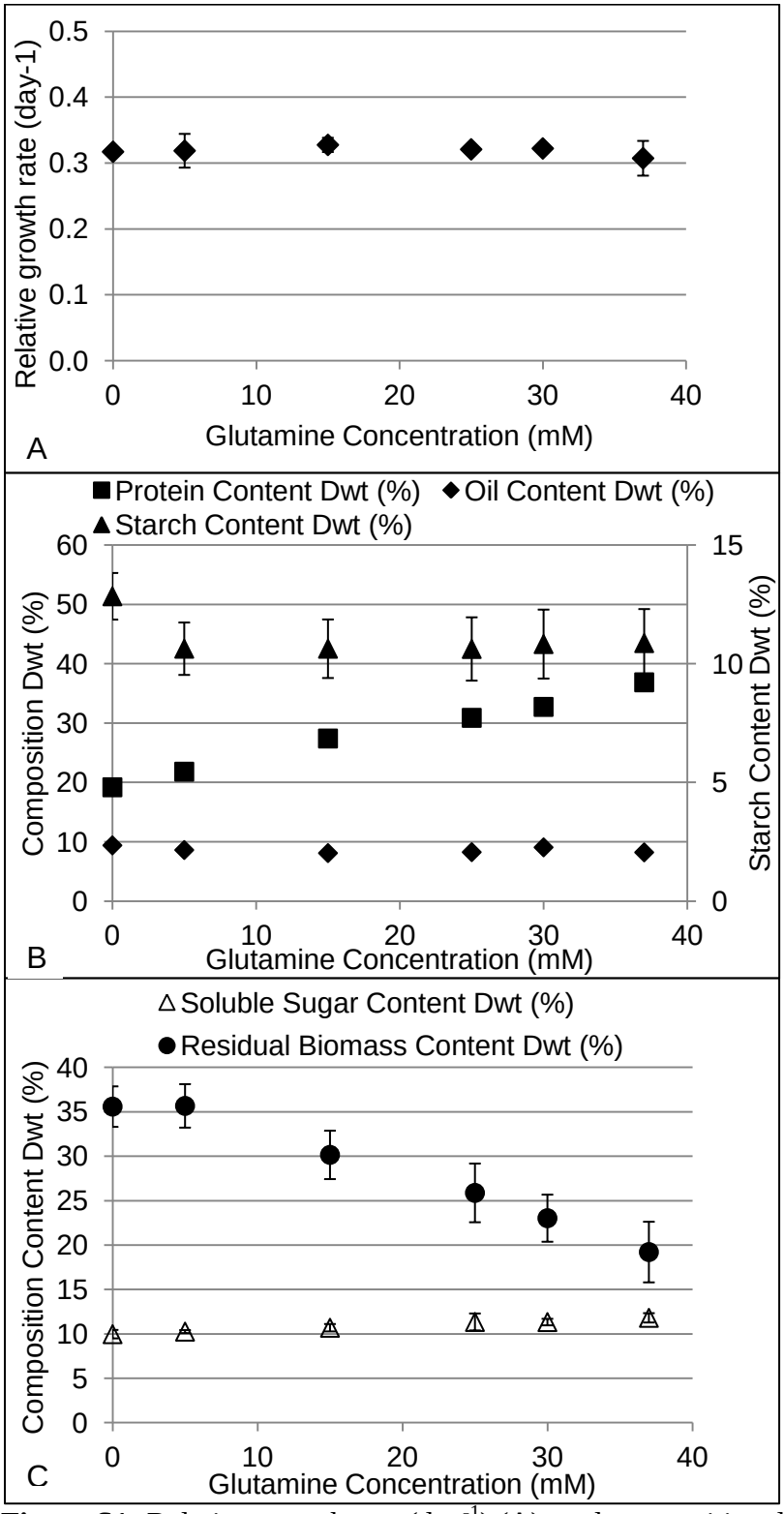


Figure S1: Relative growth rate (day^{-1}) (A); and compositional analysis of: Protein, Oil, and Starch (B); Soluble sugars and Residual biomass content (C) in soybean

somatic embryos *cv. Jack* as a function of the initial glutamine concentrations in the SHaM media; note the media was not changed during the 7 day culture period. Initial glutamine concentrations ranged from 0 to 37 mM. The initial concentration of sucrose was 88 mM.

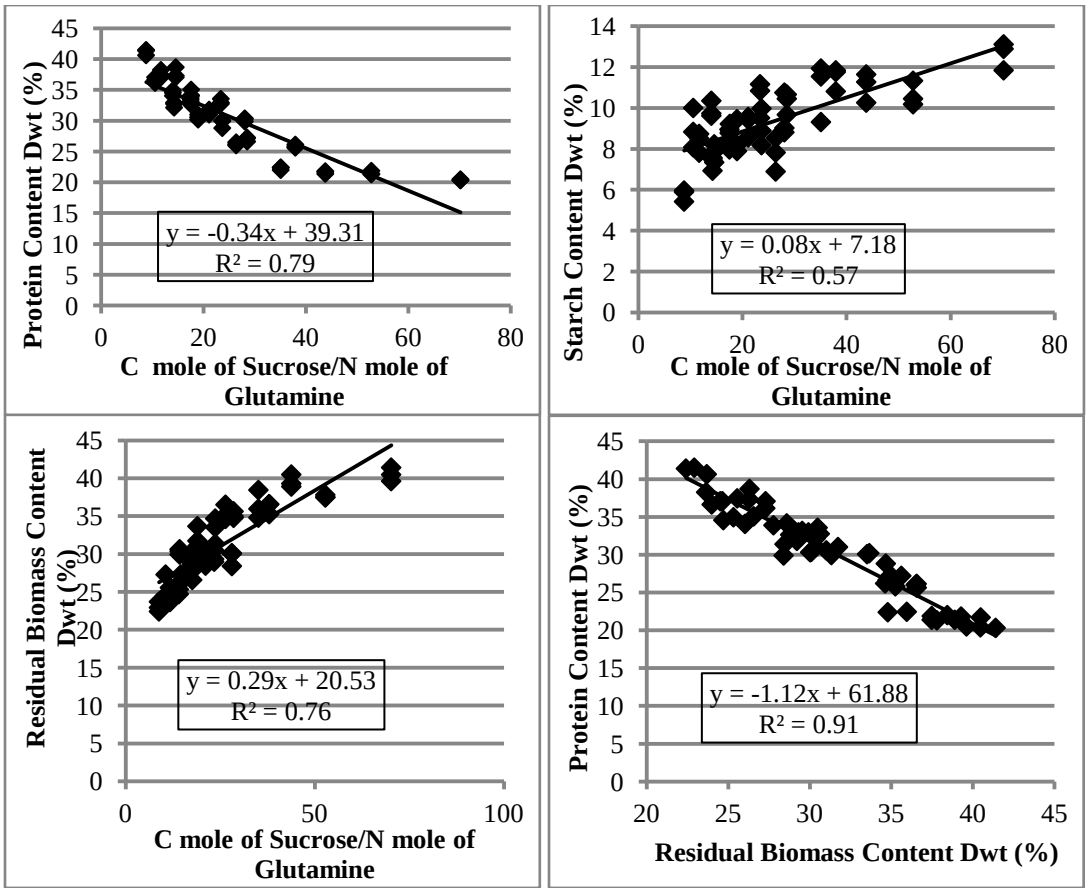


Figure S2: Linear fit correlations between media C: N ratio and the protein (A), starch (B), and residual biomass content (C) (Dwt basis) of soybean somatic embryos after 6 days in culture. The relationship between protein and residual biomass contents (D). The media in each flask was replaced after 3 days in culture.

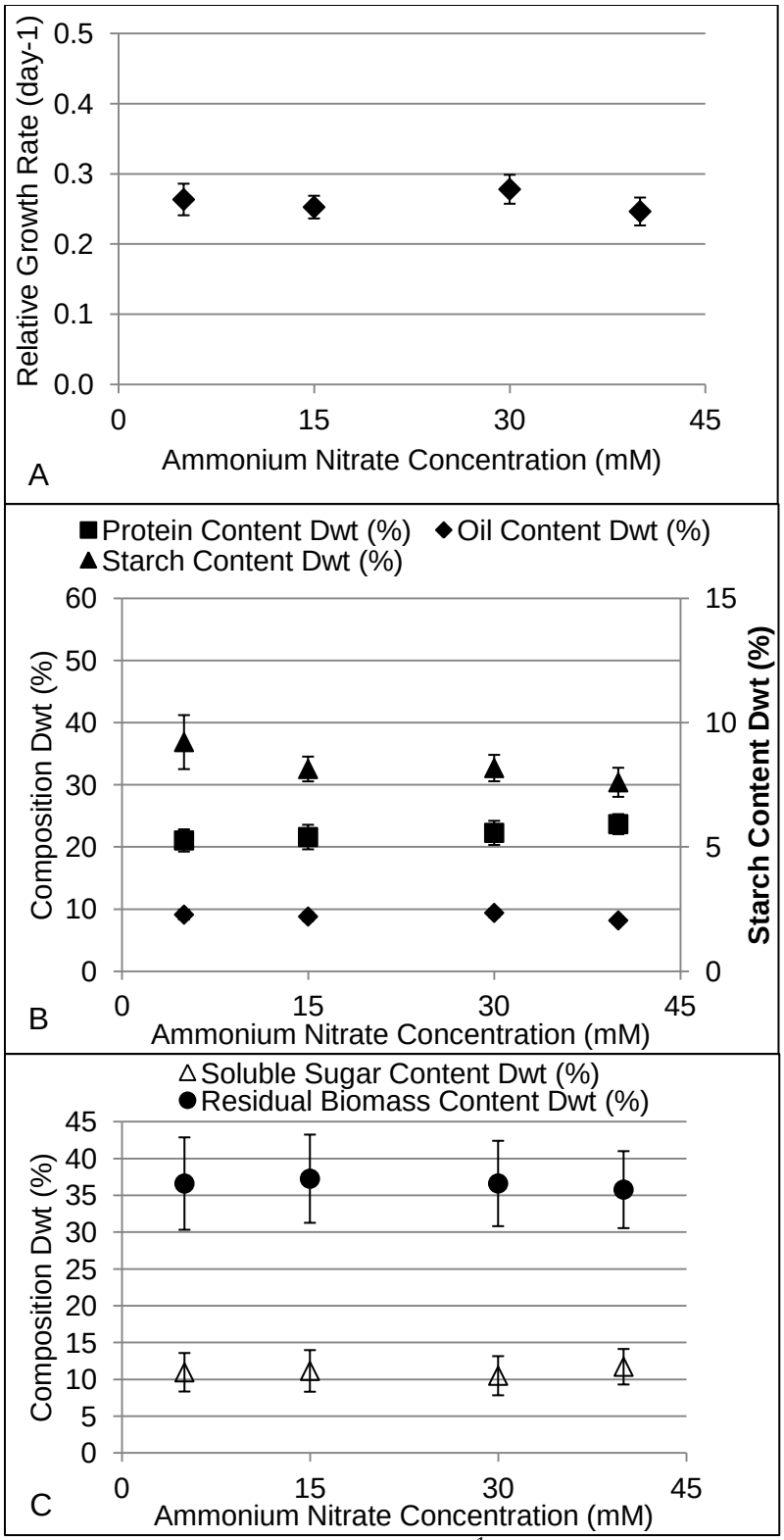


Figure S3: Relative growth rate (day^{-1}) (A); and compositional analysis of: Protein, Oil, and Starch (B); Soluble sugars and Residual biomass content (C) in soybean somatic embryos *cv. Jack* as a function of the initial ammonium nitrate concentrations in the SHaM media; note the media was not changed during the 7 day culture period.

Initial ammonium nitrate concentrations ranged from 5 to 40 mM. The initial concentration of sucrose was 88 mM.

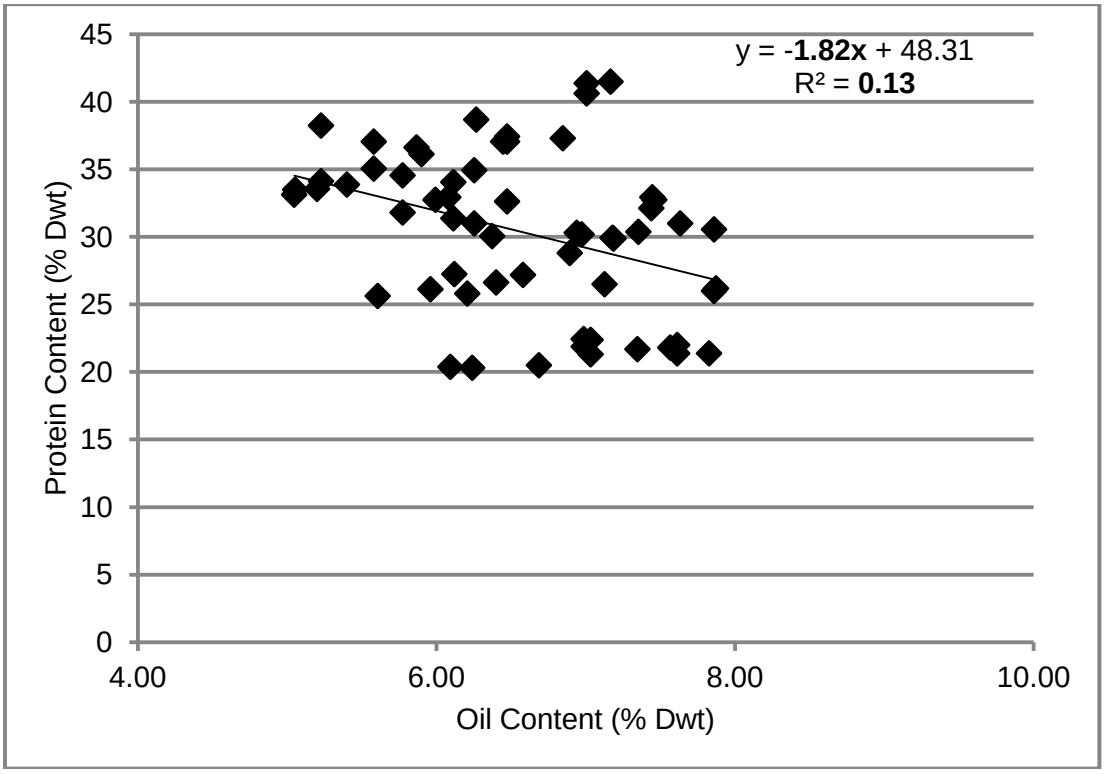


Figure S4: Correlation analysis of Oil and Protein contents (Dwt basis) of soybean somatic embryos cultured for 6 days under different initial Carbon-to-Nitrogen (C: N) mole ratios. The slope is -1.82 and linear correlation coefficient $R^2 = 0.13$. The media in each flask was replaced after 3 days of culture.

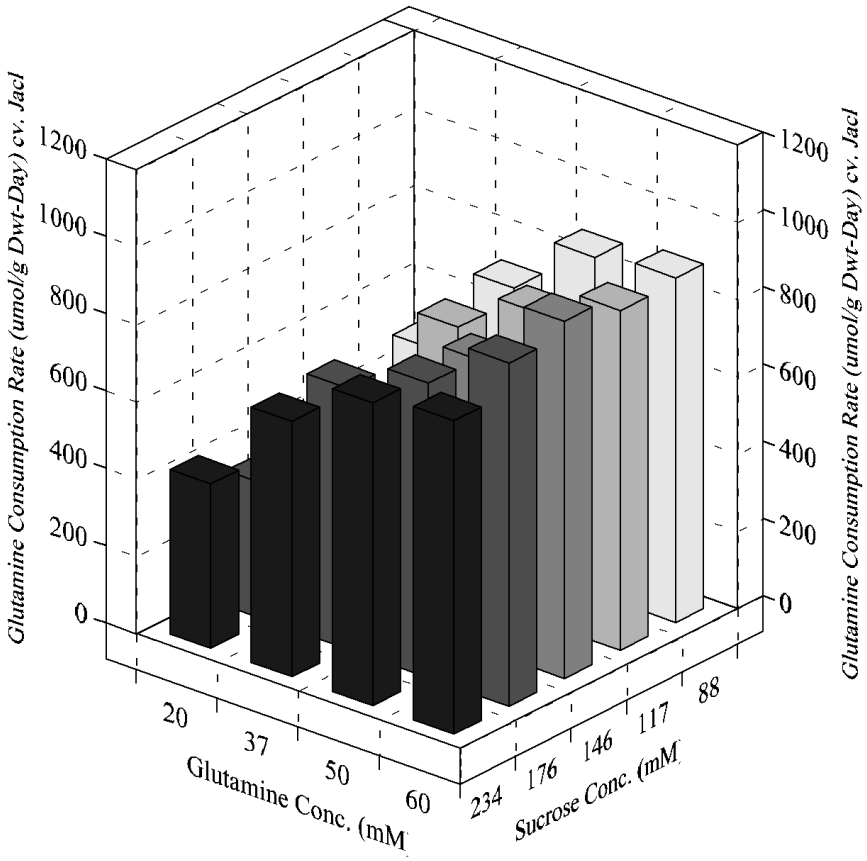


Figure S5: Comparison of the glutamine consumption rates (on dry weight (Dwt) basis (µmol $gDwt^{-1}$ Day^{-1})) of soybean somatic embryos *cv. Jack* cultured in SHaM media at a range of glutamine (20, 37, 50, and 60 mM) and sucrose concentrations (88, 117, 146, 176, and 234 mM). The media in each flask was replaced after 3 days of culture.

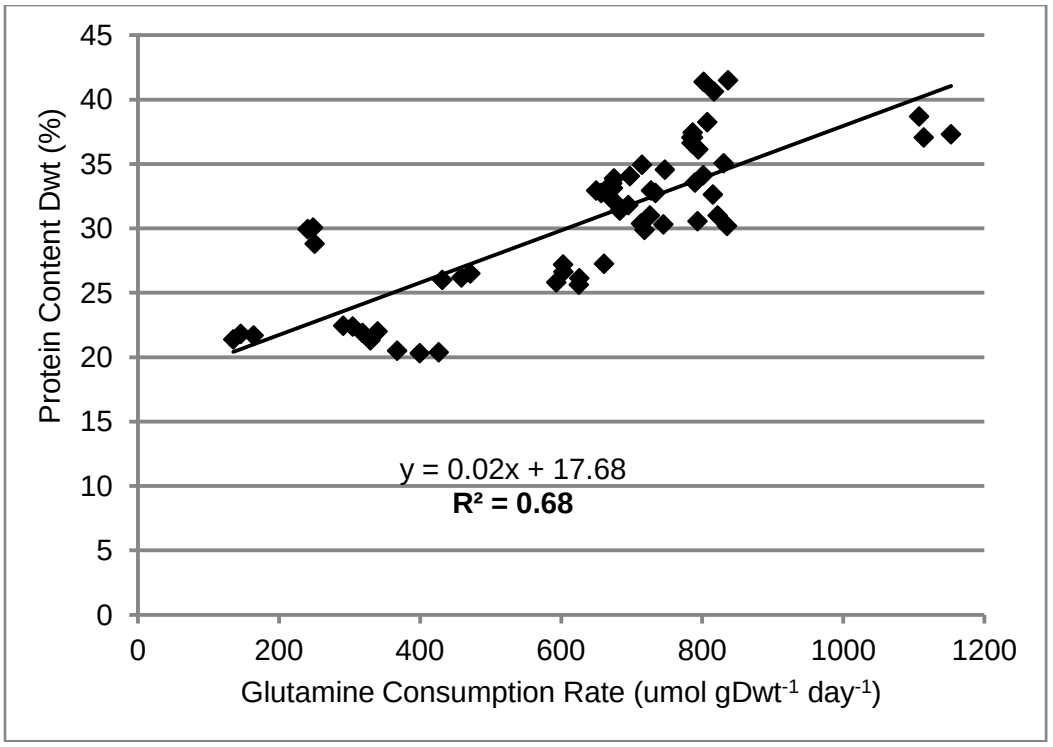


Figure S6: Linear Fit with α=0.05 of the glutamine consumption rate (µmol $gDwt^{-1}$ day^{-1}) versus the protein content of soybean somatic embryos *cv. Jack* cultured in SHaM media *in vitro* for 6 days. The media in each flask was replaced after 3 days of culture.

Table S1: Biomass composition comparison between soybean zygotic and SHaM embryos cultured in media with a range of C: N ratios.

Embryo Type	**C:N ratio**	**Protein Content Dwt (%)**	**Oil Content Dwt (%)**	**Starch Content Dwt (%)**
Zygotic Embryos	6 → 91[a]	47 → 14	Range: 9 – 14	5 → 10
SHaM Embryos	9 → 70[b]	41 → 20	Range: 5 – 8	6 → 13

Biomass composition data was extracted from the literature:
a – Allen and Young, 2013
b – This study

Reference

Allen DK, Young JD. 2013. Carbon and Nitrogen Provisions Alter the Metabolic Flux in Developing Soybean Embryos. *Plant Physiology* 161, 1458-1475.

CHAPTER 3

METABOLIC FLUX MAP COMPARISONS OF WILD-TYPE SOYBEAN (*CV. JACK*) ZYGOTIC AND SOMATIC EMBRYOS

Daft of a research paper intended for the *Journal of Experimental Botany*

Quyen Truong[1], John D. Everard[2], Jacqueline V. Shanks[1]

[1]Department of Chemical and Biological Engineering, Iowa State University, Ames, IA 50010, USA, [2]DuPont Agricultural Biotechnology Research and Development, DuPont Experimental Station, Wilmington, Delaware 19880, USA

Abstract

Soybean [*Glycine max* (L.) Merr.] seed are a primary source of vegetable protein and oil to the feed and food industries. Soybean somatic embryos cultured in Soybean Histodifferentiation and Maturation (SHaM) were examined as a model system for developing an understanding of assimilate partitioning and metabolic control points for protein and oil biosynthesis in soybean seed. This report describes the in-depth studies of zygotic and SHaM embryos under similar culture conditions and carbon and nitrogen sources. There is no significantly different in relative growth rate for both embryos. Both protein and oil content were lower for SHaM embryos than in zygotic embryos; however, starch contents were comparable, and the balance of the biomass differences, which was accounted for by the residual (structural carbohydrate) pool, was higher in the SHaM embryos. Flux analysis in cultured embryos resulted changes in nitrogen uptake and flux into oil biosynthesis, respiratory flux (CO_2), glutamine biosynthesis flux, fluxes in the total of plastic and cytosolic of triose phosphate to phosphoenolpyruvate pathway, as well as an increase in tricarboxylic acid cycle activity for zygotic embryos. However, fluxes into

structure and non-structure carbohydrates were significantly higher in SHaM embryos. Despite these differences, the NMR relative intensities of proteinogenic amino acids and labeling patterns of protein and starch-related glucosyl units were comparable between the two embryo types. Carbon labeling patterns of SHaM embryos well fitted with the metabolic network model of zygotic embryos with three compartments: cytosol, plastid, and mitochondrion. The observations described here shed light onto metabolic pathways of SHaM embryos, especially as compared to soybean seed. Our results show SHaM embryos to be a very promising model for developing seed.

1 Introduction

With global demands for increased crop yield and quality, resulting from a rapidly growing human population and a need for renewable sources of energy, green chemistry pioneering suggests that soybeans [*Glycine max* (L.) Merr.] are at the forefront of biotechnologically engineered crops. Soybeans have a very high protein content (40%), and amongst the major agricultural crops, they are uniquely high in oil (20%) (Sinclair and de Wit, 1975). For the past decade, the breeding efforts to increase soybean productivity have been very successful. However, one of the dilemmas for breeding programs is that the oil and protein contents of soybean seeds appear to be inversely related (Hartwig and Kilen, 1991) with an approximate trade-off of 1% of oil for every 2% of protein (Clemente and Cahoon, 2009). A highly desirable goal for breeding and research programs, therefore, is to increase the oil content of soybeans at the expense of other seed components (starch and soluble

carbohydrates), while maintaining protein contents at a level that meets market demands. To attain this goal, an understanding of carbon and nitrogen partitioning into the protein and oil pools is very important, and such information could enhance the chances of successfully engineering soybean seeds to improve their quality and economic value.

In recent years, metabolic analysis of cultured developing seeds has shed light on the genes and metabolic pathways involved in the regulation of storage reserve synthesis in plant model systems, including numerous agriculturally important crop species such as: soybean (Allen *et al.*, 2009; Allen and Young, 2013; Iyer *et al.*, 2008; Sriram *et al.*, 2004), rapeseed (Junker et al., 2007; Schwender and Ohlrogge, 2002; Schwender et al., 2003; Schwender et al., 2006), and *Arabidopsis thaliana* (Hay and Schwender, 2011, 2011; Lonien and Schwender, 2009; Masakapalli *et al.*, 2010). It has become possible, with information from these studies, to contemplate rational transgenic manipulations as a way to impact seed quality. However, it would be inefficient and impractical to systematically investigate the function of a large number of genes implicated in resource partitioning using a transgenic approach based on developing zygotic soybean embryos as an experimental system because of the long lifecycle of soybeans.

Soybean somatic embryos cultured in Soybean Histodifferentiation and Maturation (SHaM) media, called SHaM embryos, provide an alternative and more practical approach. They are compositionally and metabolically similar to zygotic embryos, and can provide stable transgenic cultures for experimentation directed at understanding genotype-phenotype relationships (Nishizawa *et al.*, 2010). Under

various carbon-to-nitrogen ratios, SHaM embryos display a similar trend in the biomass composition as compared to zygotic embryos (Allen and Young, 2013). Furthermore, SHaM embryos have proven to be an excellent system for the discovery of genes that lead to oil content increases in mature soybean seed (Meyer *et al.*, 2012). In addition to the application of novel gene screening systems, SHaM embryos have been employed for the validation of gene function in soybean seed (Nishizawa and Ishimoto, 2009). Perhaps the greatest advantage provided by somatic embryos is that they can be generated and examined more quickly than zygotic embryos. The cycle time from transformation-to-harvest of somatic embryos in SHaM media is approximately 8 to 10 weeks at which time stable transgenic events can be identified and maintained for experimentation. In contrast, it takes approximately 11 months to generate seed-bearing transgenic plants, and then several more generations are required to obtain homozygous plants capable of generating zygotic embryos for study. It is also possible to initiate somatic embryos from stably transformed soybean plants. Therefore, SHaM embryos appear to be a promising model system for investigating the effect of nutritional and genetic manipulations on embryo biomass composition; however, little is known about how closely SHaM embryos match zygotic embryos at the level of metabolic control, an understanding that will be essential to guide future rational transgenic manipulation.

Steady state ^{13}C-metabolic flux analysis (MFA) has emerged as a powerful set of techniques for developing a deep understanding of metabolic control at the cellular level (Kruger et al., 2012; Libourel and Shachar-Hill, 2008; O'Grady et al., 2012; Ratcliffe and Shachar-Hill, 2006; Schwender, 2008; Stephanopoulos and Vallino,

1991; Wiechert, 2001). Steady state ^{13}C-MFA utilizes the patterns of labeled metabolites, resulting from isotopic labeling of carbon experiments, as well as input and output measurements, to establish internal metabolite flux values to generate detailed flux maps. Utilizing a generic ^{13}C metabolic flux analysis tool (^{13}C MFA), NMR2Flux model was developed for zygotic soybean embryos (Sriram *et al.*, 2004) and this model was recently used to identify the impact of temperature on carbon partitioning into protein and oil in zygotic embryos (Iyer *et al.*, 2008). In this study, we have applied ^{13}C MFA to develop a better understanding of how the metabolism of soybean SHaM embryos represents that of zygotic embryos (seed). The analyses reveal: 1) the relative growth rate, water content, and starch accumulation between two embryo types are not significantly different; 2) ^{13}C labeling patterns indicate that the metabolic intermediates between two embryos are similar; 3) the glucosyl units derived from starch and protein hydrolysis fraction are well matched between zygotic and SHaM embryos; 4) the trends of flux distribution and the major carbon flow into three compartments of plastid, cytosol, and mitochondrion of SHaM embryos are comparable to zygotic embryos. The results provide insight into potential advantages of SHaM embryos as a model system for using metabolic flux analysis to understand the regulatory mechanisms controlling assimilate partitioning in soybean seed.

2 Materials and methods

2.1 *Plant material and culture conditions: zygotic embryos*

Soybeans (*Glycine max*) *cv. Jack* were grown in soil in controlled environment chambers under the following conditions: 26°C continuous temperature, and 50% relative humidity, over a 14 hour photoperiod; with high output fluorescent lamps

providing a light intensity of approximately 500 µmol photons $m^{-2}s^{-1}$. Sufficient water and nutrients were applied to the soil to ensure vigorous plant growth. Flower clusters were tagged with dated labels when they became receptive (flowers opened that day). At 21 days after flowering (DAF), pods were harvested from the central nodes. Pods were surface sterilized with 20% bleach and 70% ethanol as follows: approximately 10 pods were placed in a sterile Petri-dish containing approximately 30 ml of 20% bleach for 10 minutes. The pods were then transferred to another sterile Petri-dish that contained approximately 30 ml of 70% ethanol for another 10 minutes and were finally washed three times with filtered DDI water; each time for 5 minutes. Under aseptic conditions, the surface sterilized pods were opened and the central seeds were selected. After removing the seed coat and embryonic axis one cotyledon was placed into a 25-ml Erlenmeyer flask containing 4 ml liquid media. The other cotyledon was placed in a pre-weighed micro-centrifuge tube to gain an estimate the initial fresh weight of the cotyledon that was placed into culture. These cotyledons were then frozen in liquid nitrogen and lyophilized (see below) to attain a dry weight.

Typically, the cotyledons were between 50-80 mg fresh weight and were selected for uniform initial size. Eight replicate samples, each containing one cotyledon, were cultured aseptically for 6 days in eight separate 25-ml Erlenmeyer flasks. The medium was adapted from Obendorf et al., (1982), and the components were $MgSO_4$-$7H_2O$, CaCl-$2H_2O$, $MnSO_4$-H_2O, H_3BO_3, $ZnSO_4$-$7H_2O$, Na_2MoO_4-$2H_2O$, $CuSO_4$-$5H_2O$, KI, KCl, KH_2PO_4, $FeSO_4$-$7H_2O$, Na_2EDTA, nicotinic acid, thiamine HCl, pyridoxine HCl, m-inositol, and glycine. Further, the sucrose and glutamine

concentration in the medium were 146 mM sucrose (10% Uniform-13C labeled sucrose from Isotec, Miamisburg, OH, USA) and 37 mM glutamine, and was replaced every three days (Iyer *et al.*, 2008). An additional eight cotyledons were placed in culture media without labeled sucrose and were cultured in parallel to those containing the label. Samples of unlabeled liquid media were taken daily to determine the sucrose and glutamine consumption rates. Both sets of flasks were gently shaken at 100 rpm and maintained at 26°C under continuous light between 65 – 80 µmol photons $m^{-2}s^{-1}$ light intensity. Cotyledons were harvested after 6 days in culture, rinsed three times with approximately 10 ml of distilled sterile water, blotted dry, weighed, and frozen in liquid nitrogen prior to lyophilization at -50°C and 0.0158 mbar for 72 hours. The lyophilized cotyledons were weighed (dry weight determination). Eight labeled and nonlabeled cotyledons were pooled into their own marked container before they were finely ground using a Geno/Grinder® (SPEX SamplePrep, Metuchen NJ, USA) for further processing.

2.2 Plant material and culture conditions: SHaM embryos

Cultures of proliferative soybean somatic embryos *cv. Jack* were provided by DuPont Pioneer Hi-Bred, and were maintained, by weekly subculture, in SB196 medium. SB196 medium is a modified form of MSD20 medium (Walker and Parrott, 2001), and contained FN Lite Halides, FN Lite P, B, Mo, Murashige-Skoog (MS) sulfate, MS Fe EDTA, B5 vitamins, KNO_3, L-asparagine, and 29 mM sucrose. In addition, the SB196 medium contained a moderately high concentration of synthetic auxin, 2,4-dichlorophenoxyacetic acid (10 mg/l). Clusters of globular-stage embryos (approximately 10 mg) were used to initiate cultures by dropping them into 250-ml

Erlenmeyer flasks containing 35 ml SB228 liquid media (Schmidt *et al.*, 2005). The SB228 medium contained FN Lite macro salts, MS micro salts, B5 vitamins, $CaCl_2$, L-methionine, 88 mM sucrose, 30 mM glutamine, and 165 mM sorbitol. Embryos were induced, maintained, and matured under a light intensity of 35 – 50 µmol photons m^{-2} s^{-1}, provided by cool-white fluorescent bulbs, and a 16-h photoperiod. Cultures were maintained at 26 °C and were continuously shaken at 130 rpm for two weeks to build biomass prior to experimentation.

After two weeks in culture, working in a laminar flow hood, clusters of SHaM embryos collected from 10 flasks derived from the same subculture cycle, were separated into batches of approximately 40 to 50 randomly chosen, uniformly sized embryos, with an initial fresh weight of 1.00 g (measured and recorded to a precision of 0.01g). Three replicate samples, each containing 40 to 50 torpedo-stage embryos, were placed in 250-ml Erlenmeyer flask containing 35 ml of SHaM media with 146 mM sucrose (10% Uniform-^{13}C labeled sucrose from Isotec, Miamisburg, OH, USA) and 37 mM glutamine. The culture condition was the same as for the zygotic embryos as described above. The media in each flask was exchanged every three days. As with the zygotic embryos, a similar set of SHaM embryos with three biological replicates were cultured in non-labeling sucrose in order to compare the growth dynamics of culture with and without labeled sucrose. Samples of culture media (1 ml from each flask) were taken daily from the non-labeled cultures to monitor the consumption of sucrose and glutamine. Cultures were harvested after 6 days, rinsed three times with approximately 100 ml of distilled sterile water, blotted dry, weighed, and frozen in liquid nitrogen prior to lyophilization at -50°C and

0.0158 mbar for 72 hours. Each lyophilized culture was weighed (dry weight (Dwt) determination), then finely ground using a Geno/Grinder® (SPEX SamplePrep, Metuchen NJ, USA) prior to further processing.

2.3 Relatively Growth Rate and Normalization of flux values

The growth rates of the zygotic and SHaM embryos were normalized based on their relative linear growth rates. Growth was expressed as the dry mass added during a given growth interval, per unit of original dry mass at the beginning of the growth interval. The relative growth rate (Day^{-1}) was calculated using the following equation:

$$\mu_{relative} = \frac{\left[\frac{X_{dayn} - X_{day0}}{t}\right]}{X_{day0}}$$

where X_{dayn} is the tissue Dwt at final harvest, X_{day0} is the initial Dwt of the zygotic embryo as described above and SHaM embryos, described below, and t is the culture duration. To determine the initial Dwt of the SHaM embryo cultures, three 1.00 gram fresh weight batches of representative embryos were prepared at the same time that the experimental cultures were being initiated (see above). The embryos were placed into pre-weighed 50-ml centrifuge tubes and were frozen in liquid nitrogen prior to lyophilization, as described above. After lyophilization, the embryo Dwt was determined and used in the calculations described above.

Absolute flux values were normalized to relative fluxes, based on a sucrose uptake rate of 100 moles accumulating in the biomass. This normalization is similar to the total biomass accumulation; therefore, it is easy to compare the flux across the

phenotypes. For example, the absolute flux through the plastidic glycolytic pathway, hxi^p, was 2000 µmol g-Dwt^{-1} day^{-1}, and the corresponding sucrose intake rate for that culture was 1000 µmol g-Dwt^{-1} day^{-1}. Therefore, the relative flux through this pathway after normalization for the sucrose intake of 100 moles was [(2000/1000)*100 = 200], the flux for hxi^p would be 200.

2.4 Biomass Quantification and Fractionation

For zygotic embryos, three samples of each labeled and nonlabeled culture, approximately 100 mg of dry powdered embryo, was weighed. A similar process was used for SHaM embryos, with six samples of each labeled and nonlabeled culture weighed closely with 100 mg dry powdered embryos. Both zygotic and SHaM embryos were extracted with the same methodology.

Each sample was extracted with 1 ml of n-hexane at 40°C for 1 hour, then centrifuged at high speed (13,200 rpm) for 10 minutes at room temperature. The process was repeated five times and after each extraction, the lipid-containing solvent was pooled into a pre-weighed glass tube (approximately 4.8 ml final volume), and dried for four days in a hood at room temperature. The mass of lipids after solvent evaporation were measured gravimetrically. The remaining hexane-extracted biomass was dried and further extracted for protein in 600 µl of 200 mM phosphate buffer (pH = 7.2) containing 14 mM β-mercaptoethanol at 4°C for 20 minutes to dissolve and suspend the protein into solution. The extracts were centrifuged at high speed at 4°C for 15 minutes, and the supernatant was transferred into a micro-centrifuge tube. The extraction was performed on the pellet two more times with 400 µl of the buffer (per extraction), and the supernatants were pooled with that from the

initial extraction (approximately 1.2 ml final volume per each sample). Prior to determining the protein content, the extracted protein from three samples of labeled zygotic embryos were combined into one sample tube. A similar process was performed for labeled SHaM embryos. Protein contents of the extracts were measured by Bradford Assay (Bradford, 1976). The soluble protein extraction also removed soluble sugars from the samples. However, preliminary studies showed that soluble sugar content from protein extraction was insignificant.

Residual soluble sugars in the defatted/deproteinated-biomass pellets were extracted into 1 ml of 80% aqueous ethanol in a water bath sonicator (Fisher Scientific FS110H Ultrasonic Cleaner) at 60°C for 20 minutes, and vortexed every 5 minutes. The extraction was repeated four times and after each extraction, the solvent containing the soluble sugars was pooled into a pre-weighed glass tube (approximately 4.8 ml final volume), and dried for 3 days in a 40 °C oven. Once the samples were dry, the soluble sugar content in these fractions was measured gravimetrically. The remaining pellets from each sample were placed in 1.5 ml of DDI water, covered with foil, and autoclaved (liquid cycle at 121°C and a pressure of 15 psi) for 30 minutes prior to starch digestion and extraction. Starch was digested in 1.5 ml of 100 mM citrate buffer (pH 5.0) containing amyloglucosidase at a ratio of 0.025 mg enzyme: 1 mg of tissue dry weight (approximately 2 µl of amyloglucosidase or 0.6 units of enzyme per sample). Samples were incubated overnight in a 30°C water bath. Starch content was quantified using a glucose assay kit (Sigma, St. Louis, MO). Extracted starch samples from zygotic and SHaM embryos were frozen in -80 ºC freezer for 4 hours and lyophilized at -50 ºC and 0.018

mbar for 72 hours. Then, starch powders were dissolved in DDI water prior to acid hydrolysis.

The percentage of the residual biomass fraction was estimated by substrate of the measured sum of the biomass, i.e., protein (by combustion analysis) + oil + starch + soluble sugar + ash and an estimation of the DNA/RNA content (see below) from the original mass of tissue extracted. Ash content was determined by Thermo-Gravimetric Analysis (TGA), as described below. DNA/RNA content was assumed to be 5 percent on a Dwt basis (Stephanopoulos *et al.*, 1998).

Ash content was determined by complete combustion of the samples in a Thermo-gravimetric Analyzer (TGA 7) running the following temperature ramp program: ramp from room temperature to 100 °C and hold for 10 min; ramp to 200 °C at 10 °C/min; hold at 200 °C for 10 min; ramp to 400°C at 10°C/min; hold at 400 °C for 30 min; ramp to 700 °C at 10 °C/min; hold at 700 °C for 180 min; drop to, and hold at, room temperature until sample removal. The remaining residue in the crucible was taken as the ash content.

Sucrose and glutamine consumption rates, from the labeled cultures of both zygotic and SHaM embryos, were determined by measuring sucrose and glutamine concentrations in the culture media at the initiation of the experiments, in the media recovered at the three-day exchange, and at culture that was harvested. However, for the non-labeled cultures, for both zygotic and SHaM embryos, rates were determined by measuring the sucrose and glutamine concentrations in the culture media daily. Each media sample was filtered with a syringe filter (25mm, 0.45 µm, Nylon, XPERTEX®) prior to injection into an HPLC (Waters 1525 Binary Pump and 717^{plus}

AutoSampler fitted with an Aminex HPX-87K column, and a Waters 2414 RI detector). Samples were eluted with a water mobile phase at 50°C, at a flow rate of 0.35 ml/minute; each run was 80 minutes. By running the separation for 80 minutes, both sucrose and glutamine were detected in the same HPLC run. The HPLC run also detected sorbitol in the liquid medium. The concentration of soluble sugars from the protein extracts were analyzed as above, except the runs were 30 minutes long to detect sucrose, glucose, and fructose. A series of sucrose (15, 44, 88, 146, 176, and 234 mM), glutamine (5, 10, 20, 30, 40, and 60 mM), and sorbitol (50, 100, 150, 200 mM) standards were prepared and measured with the unknown samples. Standard curves were generated by plotting the area under the peaks versus their concentrations, and were used to compute the concentrations of the unknown samples.

Sucrose and glutamine consumption rates were calculated using the following equation:

$$\frac{dS}{dt} = \frac{\left(S_{day6} - S_{day0}\right) \times V}{\left(t_{day6} - t_{day0}\right) \times \left(X_{day6} - X_{day0}\right)}$$

where S is the media sucrose or glutamine concentration at the onset and end of the culture period, V is the volume of the media after three or six-days of culture, t is time, and X is the embryo Dwt.

The Dwt percentages of carbon, hydrogen, and nitrogen were measured by elemental analysis using the Perkin-Elmer Model 2400 Series II CHN&S Elemental analyzer (Chemical Instrument Facility, Chemistry Department at Iowa State

University). In the determinations of biomass used throughout this study, total protein content was calculated from the elemental analysis (% N x 6.25).

2.5 Protein and Starch Hydrolysis, and NMR Sample Preparation

Protein extracts of zygotic and SHaM embryos were dialyzed using dialysis cassettes (Pierce, Rockford, IL) in 10 mM phosphate buffer (pH 7.2) at 4 °C for 24 hours to remove any soluble sugar residues in the sample. The protein and starch samples were vacuum hydrolyzed in hydrolysis tubes (Pierce Endogen, Rockford, IL) using 6N constant boiling hydrochloric acid (HCl) (Pierce, Rockford, IL). The acid was added in the ratio of 6 ml HCl:10 mg protein/starch. The sample tubes were sealed and gently vortexed for 1 minute. The hydrolysis tubes containing the protein and starch samples were evacuated multiple times, purged with nitrogen to remove residual air, and re-evacuated to prevent oxidation during hydrolysis. The protein and starch samples in hydrolysis tubes were hydrolyzed for 4 hours at 140-145°C in a heat block. After hydrolysis, a Rapidvap evaporator (Labconco, Kansas City, MO) was used to evaporate the hydrochloric acid for 2 hours at 35 % speed, 40 bar pressure, and a temperature of 45°C. Each dried hydrolysate (one for protein and one for starch) was reconstituted in 2 ml of de-ionized water using a Rapidvap evaporator for 30 minutes at 70 % speed, atmospheric pressure, and room temperature. The reconstituted samples were filtered using a centrifuge and Spin-X® centrifuge tube with filter. The filters were discarded, the protein and starch sample tubes were tightly sealed with parafilm, and holes were poked through the parafilm for vaporization of water during freeze-dry. The protein and starch samples were frozen in an -80°C freezer for 24 hours and lyophilized at -50°C and 0.0158 mbar for 72

hours or until the samples were completely dry. The protein and starch samples were reconstituted in 600 µl of deuterium oxide and vortexed. Then, a small portion of the protein sample was diluted to 20-50 µl using DDI water and analyzed by HPLC for proteinogenic amino acid content. HPLC analysis was carried out at the Purdue Proteomic Facility, Purdue University. Sixteen proteinogenic amino acids can be analyzed using precolumn derivatization with 6-Aminoquinolyl-N-Hydroxysuccinimidyl carbamate (Cohen, 2000). The remainders of the protein and starch samples were reconstituted with deuterium oxide as described above; the sample pH was adjusted to 0.8-1.0 using deuterium chloride/D_20 mixture if required; this step was crucial due to pH-dependence of chemical shifts of the carbon and hydrogen in samples. A pinch (2 or 3 flakes) of 3-(Trimethylsilyl)-1-propanesulfonic acid sodium salt (TMS) internal standard flakes was added to each protein and starch sample, vortexed well, and then transferred to a clear NMR. NMR samples were submitted to NMR facility (Department of Biochemistry, Biophysics, and Molecular Biology at Iowa State University).

2.6 NMR spectroscopy and Analysis

2-Dimensional [^{1}H, ^{13}C] Heteronuclear Single Quantum Correlation (HSQC) spectroscopy of hydrolyzed protein and starch samples were performed on a Bruker Avance DRX 500 MHz spectrometer at 298 K. The HSQC spectra were acquired using a modified version of the INEPT spectra (insensitive nuclei enhanced by polarization transfer) pulse sequence (Bodenhausen and Ruben, 1980). TMS is used as an internal standard; the reference zero ppm was set by the methyl signal of TMS. The magnetic resonance frequency of ^{13}C (F1 dimension) and ^{1}H (F2 dimension) are

125.7 MHz and 499 MHz, respectively. The spectral widths along the F1 dimension and F2 dimensions are 5028.05 Hz and 5482.26 Hz, respectively. Peak aliasing was applied to minimize the sweep width in the F1 dimension. The number of complex data points collected were 900 (^{13}C) and 1024 (^{1}H). The number of scans ranged from 16 to 32.

NMR spectra were acquired using Xwinnmr (Bruker) software and analyzed with NMRview freeware (Johnson and Blevins, 1994) using the Linux operating system. NMR cross-peaks signals of singlet "s", doublet "d", and triplet "t" can be processed directly to obtain raw signal intensities using the NMRview software. Both aliphatic and aromatic carbon atoms of 16 proteinogenic amino acids in the spectra were identified in the hydrolysates of the soybean protein samples extracted from the zygotic and somatic embryos (Sriram *et al.*, 2004). Each carbon of the proteinogenic amino acids was detected by its unique ^{13}C and ^{1}H chemical shifts (Wuthrich, 1976), as well as the distinctive coupling patterns and J-coupling constants (Jcc) (Harris, 1983). The chemical shifts and J-coupling-constant (Jcc) values were obtained from Krivdin and Kalabin, 1989. Further, to quantify overlapping multiplets [(s, d1, s2, dd), (s, d1, d2, d3, dd1, dd2, dd3, qd)] on the 2D [^{1}H, ^{13}C] HSQC spectra, peak deconvolution software, developed by Sriram et al., (2004), based on a spectral model originally proposed by van Winder et al., (2001), was used. Since the peaks of the aromatic amino acids and some of the protein-associated glucosyl units were crowded, additional 2-D HSQC spectra were acquired such that when J-scaled along the ^{13}C dimension, by integrating pulse sequences reported previously (Brown, 1984; Willker *et al.*, 1997), peaks could be resolved. J-scaling increases multiplet

separation by an even integral factor J and eliminates multiplet overlap. J-scaling factors of six were employed in processing aromatic amino acid and some of glucosyl units.

2.7 Metabolic Flux Analysis: Reaction Network for Central Carbon Metabolism

The reaction network of central carbon metabolism of soybean embryos was modeled, consisting of reactions of glycolysis (GLY), oxidative pentose phosphate pathway (OxPPP), tricarboxylic acid cycle, and anaplerotic reactions. A total of 76 metabolic reactions were formed by carbon arrangements, and the model simulated isotopomers with carbon labeling in 135 metabolic pools (Supplement Table S1). The metabolic flux map in this study (Fig. 3) is based on a flux model of developing zygotic embryos (Sriram *et al.*, 2004) with extensive modification from additional measurements and transcriptomic data for both zygotic and SHaM embryos (data not shown) as listed in Supplemental 5. For clarification, in the cytosol and plastid oxidative pentose phosphate, we combined the five-carbon atom metabolites ribose 5-phosphate, ribulose-5-phosphate and xylulose 5-phosphate into one metabolite called pentose-5-phosphate (P5P). In the cytosol and plastid glycolysis, we also combined the three-carbon atom metabolites dihydroxyacetone phosphate, glyceraldehydes-3-phosphate, phosphoenolpyruvate and pyruvate into one metabolite called triose-3-phosphate (T3P). Due to similar carbon atom rearrangement and rapid equilibration of these metabolites (van Winden *et al.*, 2001), combining metabolites is supported. The NMR data derived from the glucosyl units associated with the protein and starch pools provide crucial information about fluxes through the cytosolic and plastidic glucose-6-phosphate pools, respectively. This information can be used to estimate

carbon fluxes through the GLY and OxPPP pathways in the two separate cellular compartments (Sriram *et al.*, 2007). For more detail of the reaction network representing soybean central metabolism, the reader should refer to previous work (Sriram *et al.*, 2004). Approximately 500 simulations (each time with random start values from the free fluxes) were performed for each embryo culture in order to ensure that a global optimum had been located. Significant differences between the flux values of the two embryo types were determined by the Student's t-test ($P < 0.1$), but statistical analysis of other data are based on the Student's t-test ($P < 0.05$), using the JMP software (v. 9.0.1, SAS Institute Inc. Cary NC, USA). The error in the isotopomer measurements was used to perform random Monte Carlo simulations to generate probability distributions for the flux.

3 Results

3.1 Relative Growth Rate: Zygotic and Somatic Embryos

In this study, zygotic and SHaM soybean embryos were grown in liquid culture containing ^{13}C-labeled sucrose. The concentrations of the carbon and nitrogen sources for both embryos, cultured *in vitro,* were designed to mimic the conditions experienced by developing seed on the plant, and were set at 146 mM sucrose (10% Uniform ^{13}C-sucrose) and 37 mM glutamine (Allen *et al.*, 2009; Allen and Young, 2013; Gifford and Thorne, 1985; Hsu *et al.*, 1984; Iyer *et al.*, 2008; Obendorf *et al.*, 1982; Pipolo *et al.*, 2004; Schmidt *et al.*, 2005). Because cultures of zygotic and somatic embryos are so phenotypically different, in order to make growth rate comparisons possible, they were normalized based on a relative linear growth rate (growth expressed as dry mass added during a given growth interval per unit of

original dry mass at the beginning of the growth interval), (see "Materials and Methods"). After six days of culture on ^{13}C-labeled sucrose, each zygotic embryo cotyledon had accumulated 53 ± 10 mg of dry weight (n = 8) from a starting dry weight of 17 ± 3 mg/cotyledon, which translated to a relative growth rate of 0.52 ± 0.10 per day. Similar calculations were performed for SHaM embryos. The SHaM embryos cultured for 6 days accumulated 660 ± 15 mg dry weight per flask from an initial dry weight of 170 ± 13 mg per flask (n = 3). This gives a relative growth rate for SHaM embryos of 0.64 ± 0.01 per day, (Table 1). In the t-test comparing the means ($P < 0.05$) of the zygotic and SHaM embryos indicate that the relative growth rates were not statistically different.

Replicate sets of experiments were performed with unlabeled sucrose (146 mM) as the principal carbon source. The data show that the relative growth rates were 0.52 ± 0.11 per day (n = 8) and 0.63 ± 0.03 (n = 3) for zygotic and SHaM embryos, respectively. Based on the similar statistical analysis as mention previous, the results indicate that the relative growth rates of zygotic and SHaM embryos were not significantly impacted by the presence of ^{13}C label in the sucrose. The growth rate data show that the relative growth rates of both the zygotic and SHaM embryos are comparable under both labeled and unlabeled conditions. Further, we observed that the growth of SHaM embryos has high reproducibility relative to zygotic embryos. Therefore, SHaM embryos provide an ideal model system application of metabolic flux analysis studies to test hypotheses regarding carbon partitioning.

3.2 Biomass Fractionation: Zygotic and SHaM Embryos

Measurements of biomass composition (oil, protein, soluble sugars, and starch) are extremely important to determine the intracellular fluxes in plant tissues. The percent oil content on a Dwt basis, determined by exhaustive hexane extraction for both soybean zygotic and SHaM embryos, is shown in Figure 1. The oil content of zygotic embryos under unlabeled and labeled conditions was comparable at 15.5% ± 0.3% and 15.2% ± 0.1%, respectively. In contrast, the oil content from SHaM embryos under unlabeled and labeled conditions were 8.9% ± 0.3% and 8.6% ± 0.3%, respectively. The SHaM oil content was significantly lower than the oil content measured in zygotic embryos (a 42 % difference); however, the data clearly shows that the presence of ^{13}C label had no effect on oil pool in these tissues. Further, Figure 1 shows that uniformly labeled sucrose substrate did not have any significant effect on biomass profiles of protein, starch, or soluble sugars in both SHaM and zygotic soybean embryos. Protein content was determined based on the nitrogen content in residual insoluble material (elemental analysis, see "Materials and Methods"). The total protein from SHaM embryos was 29.1% ± 0.7% compared to 37.8% ± 0.1% for zygotic embryos, a 23% difference. This confirms the reduction in oil and protein observed earlier for SHaM embryos when compared to zygotic embryos (He et al., 2011; Li et al., 2010; Schmidt et al., 2005). Further, the quantitative analysis of starch data obtained for both embryos indicates that there is no difference in the total starch extraction pool, 9.9% ± 0.5% and 9.0% ± 0.9%, for zygotic and SHaM embryos, respectively. Our starch data are in good agreement with the zygotic and SHaM embryos *in vitro* studies (Allen and Young, 2013; Iyer *et al.*, 2008). However, there was a shift in biomass composition of the cultured embryos.

For SHaM embryos, there was a large increase in soluble sugar fraction, 9.8% ± 0.3% as compared to 6.3% ± 0.7% for zygotic embryos, a 36% increase (Figure 1). To complete the biomass composition balance, the residual fraction was estimated based on mass balance for both embryos, as described in Material and Methods section. The percentage residual biomass fraction was calculated by subtracting the sum of the biomass composition [protein (by combustion analysis) + oil + starch + soluble sugar + ash + DNA/RNA (estimated)] from the mass of tissue extracted. The residual biomass fraction of SHaM embryos was significantly higher than the zygotic embryos, 26.6% ± 2.3% versus 18.2% ± 2.1%.

3.3 Sucrose and Glutamine Consumption Rates

The dominant carbon and nitrogen sources supplied to soybean zygotic and SHaM embryos *in vitro* studies have been reported to be sucrose and glutamine (Gifford and Thorne, 1985; Hsu et al., 1984; Pipolo et al., 2004; Saravitz and Raper, 1995; Schmidt et al., 2005). The sucrose and glutamine uptake rates were calculated by measuring the depletion of both substrates from the medium at day 0, day 3, and day 6 of culture (see Materials and Methods). The glutamine consumption rate follows the same trend that was observed in the protein content of zygotic and SHaM embryos. As mentioned earlier, the zygotic embryos were cultured in individual flasks with 4 ml of media and the media was replaced after three days. The glutamine accumulation rate for the zygotic embryos was 750 ± 25 μmol $gDwt^{-1}day^{-1}$. For SHaM embryos cultured in a flask of 35 ml of liquid media, the media was replaced after three days, and the glutamine accumulation rate was 490 ± 20 μmol $gDwt^{-1}$ day^{-1}, as shown in Table 1. To validate the method of calculated substrates consumption

rate, an independent experiment was set to collect sample of media and culture daily for the polynomial fit and the end point calculation for day 0, 3, and 6 (see Supplement 1). The result indicated that there was no significant difference between both methods. Therefore, all of the calculations of glutamine and sucrose uptake rates were based on the end-point calculation. As mentioned above, the protein content for the zygotic embryos was 20% higher than SHaM embryos. Zygotic embryo cultures consumed more reduced nitrogen than do SHaM embryos cultures (Table 1); and therefore, produced higher protein content, consistent for the role of reduced nitrogen source for protein biosynthesis in soybean embryos (Allen and Young, 2013; Pipolo *et al.*, 2004; Saravitz and Raper, 1995). The same observation can be made for the sucrose uptake rates of the two embryo types. The sucrose consumption rate was significantly higher in zygotic embryos than in SHaM embryos, 1360 ± 150 as compared to 990 ± 50 μmol $gDwt^{-1}day^{-1}$, as shown in Table 1. The role of sucrose as a main carbon source in the cultures for the two embryo types was more complex. Therefore, we might expect higher carbon flow through metabolic pathways in zygotic embryos relative to SHaM embryos.

3.4 Proteinogenic Amino Acid and Glucosyl Units

Protein samples of zygotic and SHaM embryos were hydrolyzed into the corresponding amino acids. The amino acid profiles of zygotic and SHaM embryos were compared as shown in Figure S1. Although the total protein content was significantly higher for zygotic embryos than for SHaM embryos, the overall composition of the amino acid, which was bound to protein, did not change, with the exception of glutamate/glutamine (Figure S1). Due to sample limitation, each culture

was analyzed only once; therefore, no error bars are presented. However, we are confident that our data did not have large variation because the analysis of proteinogenic amino acid of SHaM embryos was repeated in our laboratory using a similar procedure to the Purdue Proteomic Facility (see Materials and Methods), with three instrument replicates. Figure S1 shows the accuracy level from different instrumentation for proteinogenic amino acid profiles with Purdue University versus our laboratory HPLP analysis.

Furthermore, the NMR relative intensity of each amino acid, for the zygotic and SHaM embryos, were compared to provide detailed differences, as shown in Figure 3. When the 2D [^{1}H, ^{13}C] NMR relative intensities of each amino acid were plotted against each other, they produced a line with a linear correlation coefficient of R^2 = 0.90 and a slope of 0.96. This data provides a clear indication that the ^{13}C labeling patterns of SHaM and zygotic embryos were related, but that there were physiological differences between the two. Furthermore, the slope value of the two embryos showed that the metabolic intermediates between the two embryos were similar. A comparable observation was found when zygotic embryos were compared to different cultivar of the exact media and growth condition. Results showed a linear correlation coefficient R^2 = 0.96 and a slope value of 0.97 (unpublished data), therefore indicating that the carbon-labeled position on the zygotic embryo was the same with different cultivar.

Another relevant observation comes from a correlation between zygotic and SHaM embryos for the glucosyl units (levulinic acid (LVA) and hydroxyacetone (HyA)) derived from the hydrolysate of protein and starch fraction. Figure 4 shows

the relative 2-Dimensional [^{1}H, ^{13}C] HSQC spectrum multiplet intensities of glucosyl units derived from starch and hydrolyzed protein from both embryos. The relative intensities of labeling in the glucosyl units from protein and starch hydrolysis were well matched between zygotic and SHaM embryos. The similar trends in those glucosyl units of protein and starch hydrolyzed were the same between zygotic and somatic embryos, which indicated that the distributions of fluxes in the glucose-6-phosphate pools of the plastid and cytosol compartments between the two embryo types are highly similar. Table S1 lists each amino acid and glucosyl unit with its corresponding to its precursor, as well as its isotopomer position.

3.5 ^{13}C-Metabolic Flux Analysis and Steady State Verification

A mathematical framework of ^{13}C Metabolic Flux Analysis (^{13}C-MFA), NMR2Flux, for metabolic network used here is based on former studies of developing soybean embryos (Sriram *et al.*, 2004), and has successfully explained various aspects of central carbon metabolism (Iyer *et al.*, 2008). The metabolic network represents three compartments: cytosol, plastid, and mitochondrion. Based on labeling data available for the glucosyl units on the protein (derived from cytosolic hexose phosphate) and starch (synthesized from plastic hexose phosphate) hydrolysates (Sriram *et al.*, 2007), the model represents parallel pathways of glycolysis and pentose phosphate for compartmentalization into the cytosol and the plastid. In both embryo phenotypes, we verified that the isotopomeric compositions of the hexose pools of cytosol and plastid compartment were different. The results were obtained by comparing the LVA and HMF peaks of protein and starch hydrolysates of zygotic and SHaM embryos, as shown in Table S1. The data

indicated that most of the abundances are statistically different. Clearly, this suggests the presence of separate and parallel pathways of glycolysis and pentose phosphate pathway in both cytosol and plastid of both zygotic and SHaM embryos. The isotopomeric compositions of the cytosolic and plastidic triose phosphates were not significantly different, based on comparisons of the multiplet intensities of Ala α, Phe α, and Tyr α. The plant biochemistry of soybean indicates that Phe and Tyr synthesize from plastidic PEP and hence, reflect the isotopomer composition of the plastidic triose phosphate. Ala is synthesized in cytosol and plastid (http://pmn.plantcyc.org), and therefore the composition displays a combination of isotopomeric composition in the triose phosphates from both compartments. Figure S3 illustrates that corresponding isotopomer abundances are similar for both zygotic and SHaM embryos. Thus, the data suggests that the flux values of gap, eno, pyk, in cytosol and plastid compartments were indistinguishable.

To further verify the network topology of our soybean system, in-depth computational models were performed using the zygotic embryos with 10% U-^{13}C sucrose. We compared three metabolite models that could best account for the labeling patterns. These models differenced in the locations of cytosol and plastid compartments for pentose phosphate pathways (PPP): (1) our previous model contained the oxidative and non-oxidative reactions of PPP for the plastid and the cytosol; (2) the same model as (1) with addition of Rubisco bypass in the plastid; (3) contained oxidative and non-oxidative reactions in the cytosol and non-oxidative with Rubisco bypass reactions in the plastid as shown in Supplement 3. We observed that the first two models indicated that the data equally well fitted with the statistic chi-

square (χ^2) values, 140 ± 30 and 130 ± 20, respectively, were insignificant. However, the third model displayed that χ^2 was 230 ± 430, which illustrates that the fluxes in that model were unidentifiable. The results from the second and third metabolic models show that the flux through Rubisco was very small (see Supplemental 3). We further investigate into the isotopomer balance of Rubisco pathway to track the assimilation of internal carbon rearrangement of Rubisco reaction as shown in Supplemental 4. Through the isotopomer balance in 3-Phosphoglycerate (3PG) formation, as a function of glycolytic flux (GAP) and the flux through Rubisco (P5P and CO2), it was revealed that the fluxes through the Rubisco bypass reaction (*rubisco*, Table S2) are presented to be small values for both embryos (Supplemental 4). This illustrates that the hexose breakdown in the plastid of glycolysis is independent of the Rubisco bypass for soybean embryos (Schwender *et al.*, 2004).

The major assumption for the ^{13}C-MFA model is for isotopic steady state conditions of the labeling studies of soybean zygotic and SHaM embryos. The steady state flux analysis implies that intracellular metabolite pools do not change over the time span during which the labeling experiment takes place. Three measurements can be obtained to validate the metabolic steady state condition: sucrose consumption rates, protein and starch accumulation, and growth characteristic (Kruger *et al.*, 2012). The developing study of linseed embryos indicates that the isotopic steady state was determined to be about 18 hours (Troufflard *et al.*, 2007). The zygotic and SHaM embryos were cultured in the sucrose-labeled liquid media for 6 days, showing that the isotopic steady state is attained within eight residence times. Therefore, for both embryos, the metabolite pools are expected to be constant. Furthermore, the

core of this study is based on the analysis of the ^{13}C-experiment labeling of storage products of protein and starch. The products of protein and starch hydrolysate extracted are biosynthesis-derived from the central metabolites. As a result, the labeling in proteinogenic amino acids and glucosyl units represents the labeling fingerprint of the central metabolites during protein and starch biosynthesis.

A comparison of experimental and simulated data of carbon labeling patterns indicated a strong fitted for zygotic and SHaM embryos with relative coefficients of 0.96 and 0.97, respectively, as shown in Figure S3. The deviation errors of experimental versus simulated for both embryo types were mainly collected from the NMR-derived abundances, which are on the order of approximately 1%. Therefore, the statistical agreement between experimented and simulated data supported the accuracy of the model. Furthermore, the network model resulted in low statistical criteria of the chi-square values (below the measurement errors) of 87 ± 25 and 95 ± 30 for zygotic and SHaM embryos, respectively. There is no significant different between the chi-square values for both embryo types concluded the student t's means comparison with ($P < 0.05$). Hence, the metabolic network model for soy SHaM embryos can be related to the developing zygotic embryos.

For comparison of flux results between zygotic and SHaM embryos, all flux values are given in both absolute (Table S2) and relative units (Table S3). The absolute units, expressed in terms of the rate of particular metabolite into the biomass, are μmol metabolite $gDwt^{-1}Day^{-1}$. For the relative unit of flux, values were normalized in terms of the uptake flux of the main carbon source of sucrose (100 moles) (Iyer *et al.*, 2008). By normalization relative to the sucrose uptake rate, a flux

comparison between zygotic and SHaM embryos on equal number of the carbon flow source of sucrose can be made. Examination of the relative fluxes and flux ratios of various nodes in the reaction network enhance the understanding of the metabolic effect on the zygotic and SHaM embryos. As both embryos are phenotypically different, this understanding could lead to an understanding of embryos without the influence of the carbon source (Stephanopoulos and Vallino, 1991). From this point on, all references to flux and flux ratios are in terms of relative flux values for the comparison between zygotic and SHaM embryos. Another normalization method that is often used in the literature is based on the constant rate of carbon flow into the biomass (Lonien and Schwender, 2009; Masakapalli *et al.*, 2010). Furthermore, the mentioned authors indicate that when both of these normalization methods were applied, the conclusions did not deviate from one another.

3.6 Metabolic Flux Map Comparison of zygotic and SHaM embryos

To determine how closely the metabolism of SHaM embryos mirrored that of developing zygotic embryos, a mathematical model of flux analysis was employed to track carbon flow through metabolite pools of the three compartments in the given stoichiometric network. Table 2 summarizes the significantly different fluxes between zygotic and SHaM embryos. Some common trends observed between both embryo types are listed below. A detailed analysis of flux on both embryos indicates no difference in the parallel pathways of glycolysis and OxPPP between the cytosol and plastid, based on the labeling data. The flux through non-oxidative reactions of the OxPPP catalyzed by transketolase (tkt) and transaldolase (tal) were greater in plastid than in the cytosol, see Table S3. In terms of flux distribution, the major fluxes are

phosphogluconate decarboxylating (pgl), 3-phosphoglyceraldehyde (gap), phosphopyruvate hydratase (eno), pyruvate kinase, and respiration. A similar trend was perceived from a different soybean zygotic cultivar at the same media and growth condition as data were compared in Table S4 (Sriram *et al.*, 2004). Furthermore, in both embryos, the hexose isomerase fluxes were in the direction of fructose-6-phosphate (F6P) to glucose-6-phosphate (G6P), and directed carbon flow through the OxPPP in a cyclic manner. Figure 2 displays a flux map of SHaM as compared to zygotic embryos. The thickness of the arrow indicates the difference of magnitude of the fluxes. When analyzing the relative flux, moles of metabolites per mole of sucrose consumed, with the total flow of carbon into biomass kept constant across both embryos, the following fluxes stayed the same: oxidative and non-oxidative pentose phosphate pathways; gogat reaction; all amino acid biosynthesis fluxes, with the exception of glutamine; the glucose 6-phosphate, pentose 5-phosphate, and triose 3-phosphate translocators; and the plastidic malic enzyme as shown in Figure 2. Therefore, the rate of carbon flow in these pathways for zygotic and SHaM embryos are equivalent.

Investigating the carbon flow differences between the two embryos is important for understanding the partitioning of protein and oil. For both embryos, there are significant fluxes through glycolysis from triose to pyruvate (plastid pyruvate dehydrogenase was 60 to 65% of total flux to pyruvate (pyk)). This demonstrates that sugars are an important carbon source for the fatty acid metabolism. Moreover, the glycolysis fluxes from triose to pyruvate (gap, eno, and pyk) were statistically higher (approximately 25% to 38%) for zygotic embryos than for SHaM embryos (Table 2).

Further, the plastid pyruvate dehydrogenase (pdh^p) was 37% higher in the zygotic relative to SHaM embryos. However, the mitochondrial pyruvate dehydrogenase (pdh^m) was only 9% higher in zygotic than in SHaM embryos. Together, the glycolysis and pyruvate dehydrogenase produce one energy ATP and two reducing equivalents that are central for the incorporation of an acetyl-CoA group for fatty acid synthesis. Therefore, SHaM embryos have lower oil content than zygotic embryos. The anaplerotic reactions and malate shuttle are crucial parts of the interaction, and act as gatekeepers in carbon partitioning between the three compartments of soybean zygotic and SHaM embryos. Cytosolic phosphoenolpyruvate carboxylase (ppc) enzyme of anaplerotic reaction and the malate transporters (MalT1 and MalT2) assist in the flux exchange between the three compartments of cytosol, mitochondrion, and plastid (Figure 2). The flux through the ppc enzyme was lower in SHaM embryos than in zygotic embryos by 38%, which effectively could reverse the net flux through the MalT1 transporter relative to zygotic embryos. The reaction of ppc was converted to malate, which will have a dual function, MalT1 transported flux from cytosol into the mitochondrion for TCA cycle and MalT2 transported flux into the plastid for the synthesis of protein and oil. The MalT1 flux was 80% lower in SHaM embryos, but the standard deviation of MalT1 flux in SHaM embryos was large, indicating that this flux is not identifiable. Further, the result indicates that MalT2 was higher in zygotic as compared to SHaM embryos, with a difference of 50%. However, the reversibility of MalT1 for both embryos was very high (> 98%) showing that there was a quick exchange of flux between the two pools. Furthermore, the following fluxes are higher in zygotic than SHaM embryos: plastidic phosphofructokinase reaction, TCA

cycle, triose reactions (gap, eno, pyk), and malate dehydrogenase reaction as shown in Figure 2.

The differences in oil, protein, soluble sugars, residual biomass, and glutamine consumption rate between zygotic and SHaM embryos are incorporated into the changes of the biomass fluxes is listed in Table 2. The estimated CO_2 (respiration flux value from NMR2Flux simulated flux value) evolving from the oxidative pentose phosphate pathway, TCA cycle and pyruvate reactions was less in SHaM embryos than in zygotic embryos, with a 26% difference. Furthermore, mass balances of carbon between the two embryo types were generated. The main input carbon sources were sucrose and glutamine in the liquid media and the products were CO_2, lipid, protein, starch, and sugars. The result of this mass balance analysis indicated the carbon efficiency was not statistically different, approximately 85 ± 5%, for both embryo types. As expected, the fluxes in biosynthesis of starch, polysaccharide, and nucleotides (bios 4, 5, and bios 7) were higher in SHaM relative to zygotic embryos. The results agree with the higher values of biomass composition of soluble sugars and residual biomass (mainly cell wall and nucleotides) in the SHaM embryos than in zygotic.

4 Discussion

This study was designed to investigate the use of a soy SHaM embryo system as a physiological model of seed development using steady-state ^{13}C Metabolic Flux Analysis (^{13}C MFA), particularly as a model system for understanding fundamental aspects of controlling resource partitioning. SHaM and zygotic embryos were cultured in media using the same carbon and nitrogen sources. To our knowledge, this

is the first reported study of flux models on SHaM embryos in comparison to zygotic embryos. ^{13}C MFA is able to identify compartmentalization in a plant system (Masakapalli *et al.*, 2010; Sriram *et al.*, 2004) and quantification of the response to environmental perturbations (Allen *et al.*, 2009; Allen and Young, 2013; Iyer *et al.*, 2008) or genetic engineering (Alonso *et al.*, 2007; Lonien and Schwender, 2009). The information from our study provides a numerical comparison of relative growth rate, biomass composition, and central metabolite fluxes on SHaM and zygotic embryos. The isotopic steady-state condition of ^{13}C MFA was verified for both soybean embryo types. The isotopic steady-state was based on what was obtained for linseed embryos case study, which indicated that the steady-state time was 18 hours (Troufflard *et al.*, 2007). Both zygotic and SHaM embryos were cultured in labeling carbon source for 6 days. Therefore, the metabolic flux analysis presented in this study analysis is satisfactory under the major assumption of isotopic steady state.

^{13}C MFA is a powerful tool to track carbon flow within the plant system, to ultimately elucidate physiological changes. Due to the nature of the plant system, which includes compartmentalization of metabolic pathways, the measurement of isotope labeling patterns could lead to poor identifiable fluxes with low statistical confidences (Nargund and Sriram, 2013). To improve measurement of patterns depending on specific fluxes in the compartment of interest, multiple labeled substrates can be used to enhance constrain the metabolic network (Masakapalli et al., 2010; Nargund and Sriram, 2013). Even with experiments using multiple labeled substrates (U-^{13}C, 1-^{13}C, 2-^{13}C of glucose), Masakapalli and his co-workers showed isotopic labeling patterns could not distinguish the three different metabolic network

models. Furthermore, computational designs indicated that 1,2-^{13}C glucose alone or a combination of 3-^{13}C, 1-^{13}C and U-^{13}C glucose, can provide an effective way to the plant pentose phosphate pathways in the plastidic and cytosolic compartments (Nargund and Sriram, 2013). However, these labeling substrates are limited by commercial availability, cost, and difficulty in evaluating labeling patterns. Although in our study we used one tracer of U-^{13}C Sucrose, our soy flux model incorporated transcript profiling data from both zygotic and SHaM embryos (data not shown) to refine and verify the metabolic network topology. Therefore, major metabolic fluxes in the SHaM system should be considered relevant, when compared to seed development in-plant, and important for the future understanding of carbon partitioning with respect to media composition and transgenic manipulation.

The growth rate results for zygotic and SHaM embryos should be compared with respect to relative growth rates, because both culture embryos being phenotypically different. The relative growth rate of SHaM embryos was not statistically different from zygotic embryos, 0.62 ± 0.03 (day^{-1}) and 0.52 ± 0.10 (day^{-1}), respectively, under similar sucrose and glutamine concentrations, 146 mM (10% Uniform-13C labeled sucrose, U-$^{13}C_{12}$) and 37 mM, respectively. Furthermore, the SHaM and zygotic embryos relative growth rates were comparable to those of three different cultivars of zygotic embryos: *cv. Evans* with a value of 0.45 ± 0.07 (day^{-1}) (cultured *in* vitro146 mM sucrose with 10% U-$^{13}C_{12}$, 37 mM glutamine) (Iyer *et al.*, 2008), *cv. Amsoy* with a value of 0.43 ± 0.04 (day^{-1}) (cultured *in vitro* 140 mM sucrose, 70 mM U-$^{13}C_6$, 35 mM U-$^{13}C_5$, 12.5 mM Asparagine) (Allen *et al.*, 2009), and *cv. Jack* (cultured *in vitro* at C:N ratios from 6 to 91 of nearly 40 individual cultures) with an average value of

0.67 ± 0.10 (Day^{-1}) (Allen and Young, 2013). Therefore, the observations on flux distribution in SHaM embryo culture should be relevant for the development of zygotic embryos. Biomass composition and transcript profiling data of SHaM embryos (data not shown) represent zygotic embryos in the early to mid-storage phase of development.

The biomass composition based on dry weight (% Dwt) of each component (protein, oil, soluble sugars, starch, and residual biomass) was found to be comparable between the zygotic and SHaM embryos, as shown in Figure 1. For cultured zygotic embryos, a high value of protein synthesis corresponds to the increased uptake of nitrogen and carbon sources, glutamine and sucrose, respectively. The data supports the current understanding of sucrose and glutamine relation to signaling molecules to trigger storage processes in plant sink tissues (Haga and Sodek, 1987; Koch, 2004; Koch, 1996; Smeekens, 2000; Thompson et al., 1977). Our data clearly show that the glutamine consumption rate was significantly different between zygotic and SHaM embryos (Table 1); therefore, zygotic embryos have a higher capacity to synthesize protein. Similar results have been reported for soy zygotic embryos *in vitro* culture studies, where protein synthesis accounted for 60 to 90% of reduced nitrogen as coming from glutamine based on the labeling experiment of glutamine (Allen *et al.*, 2009; Allen and Young, 2013). Furthermore, as was observed in development soybean seed, for the process of protein biosynthesis in soybean, the metabolism of reduced nitrogen and inter-conversion of amino acids required organic acid to be mostly derived from sucrose as carbon skeleton (Pandurangan *et al.*, 2012). The protein and oil contents reported here are consistent

with other embryo cultures, and have been used to study the metabolism of developing soybean zygotic embryos (Allen *et al.*, 2009; Allen and Young, 2013; Iyer *et al.*, 2008; Pipolo *et al.*, 2004; Saravitz and Raper, 1995; Sriram *et al.*, 2004) and SHaM embryos (He *et al.*, 2011; Li *et al.*, 2010; Schmidt *et al.*, 2005). In SHaM embryos, the oil content was 50% less than zygotic embryos. This discrepancy may represent a difference in oil accumulation during the storage phase in developing soy zygotic and SHaM embryos. Our current understanding of soybean major storage biosynthesis indicated that soy oil should accumulate linearly, until reaching maturity (Rolletschek *et al.*, 2005). As for SHaM embryos, the reduced observed oil content could be due to the limitation of oil biosynthesis at the cotyledon development stage. Even though the oil content of SHaM embryos was lower than zygotic embryos, SHaM embryos have been shown to be an excellent screening tool for transgenic strategies to increase the oil content in mature seed (Meyer *et al.*, 2012). Moreover, SHaM embryos have been used for quick validation of gene function in the synthesis of seed components (Nishizawa et al., 2010). Most importantly, for the biomass composition, SHaM embryos respond to the perturbation of media treatment (data not shown) with a similar trend to zygotic embryos (Allen and Young, 2013).

The carbon labeling patterns of glucosyl units of starch and protein hydrolysate data provide related evidence of physiological similarities between zygotic and SHaM embryos. As observed in the relative intensity of the NMR analysis, LVA and HyA units of the glucosyl units of the starch and protein hydrolysated were identical between two embryo types, as shown in Figure 3. The result suggests that the pools of glucose-6-phosphate (G-6-P) between the cytosolic and plastidic compartments of

SHaM embryos were regulated in a similar manner to that of zygotic embryos. The result provides evidence for the metabolic role of starch accumulation of SHaM embryos. As observed in the carbon-to-nitrogen media treatment for SHaM embryos (data not shown), the starch accumulation pattern we have observed is consistent with that seen in soy zygotic embryos under a similar media treatment (Allen and Young, 2013). Furthermore, the level of starch accumulation was insignificantly different between the two embryo types. However, the soluble sugars and residual biomass contents were higher in SHaM than in zygotic embryos. One possible explanation for this is that soluble sugar, in SHaM embryos, is likely to decrease during the course of development. Soluble carbohydrates tend to accumulate early in development for both somatic and seed embryos (Chanprame et al., 1998).

The efficiency in converting carbon and nitrogen sources to biomass components for SHaM and zygotic embryos are comparable. The mass balance of the sucrose and glutamine metabolized to macromolecules and soluble compounds displays a carbon use efficiency of 85 ± 5% for both embryo types, consistent with the previous *in vitro* studies of developing soybean embryos of 85 – 87% (Allen *et al.*, 2009; Allen and Young, 2013). The estimates of carbon use efficiency for soybean embryos are dramatically higher when compared to other plant systems: including rose periwinkle hairy root culture of 24% (Sriram *et al.*, 2007); maize root tips of 42 – 47% (Alonso *et al.*, 2007); Arabidopsis cell culture of 60 – 65% (Masakapalli *et al.*, 2010; Williams *et al.*, 2008); developing sunflower embryos of 50% (Alonso *et al.*, 2007); and a tomato cell culture of 52 – 68% (Rontein *et al.*, 2002). In our culture media, there is evidence of an optimal supply of carbon and nitrogen sources for growth of SHaM

and zygotic embryos. Both embryo types must utilize a carbon source (sucrose and glutamine) to provide the energy and reducing power for the biosynthesis of macromolecules and soluble compounds.

In addition to sucrose as a major contribution of carbon skeleton, fixation of CO_2 by ribulose-bisphosphate-carboxylase (Rubisco) can make a significant contribution to the accouterment of organic acids via phosphoglycerate (3PG) for rapeseed study (Schwender *et al.*, 2004). However, the fluxes through this pathway, for both of our soybean embryos are very small. Based on our examination of the isotopomer experimental abundance with the uniformly labeled sucrose, we expected that the flux through this pathway is insignificant. Details of the small contribution of Rubsico bypass flux have been provided in Supplemental 4. Furthermore, our observed data of Rubisco bypass reaction agree with the proteomic study, which indicated that the Rubisco enzyme is expressed at a very low level in soybean but significantly high in rapeseed (Agrawal *et al.*, 2008). The authors suggested that this could due to low carbon contribution, in which the oil in soybean was less than rapeseed. Even though there is evidence of Rubisco presence in green oilseeds (Ruuska *et al.*, 2004), the insignificant contribution of Rubisco flux for both embryo types tested here could be due to enrichment nutrition of sucrose and glutamine concentrations provided in the liquid media. For SHaM and zygotic embryo experiments, the media was replaced every three days to ensure that nutrient supplies were constant for sucrose and glutamine sources throughout the experimental period. Another possibility to better identify this Rubisco pathway lay within the design of isotopomer label experiments could be the addition of tracer(s) to the U-^{13}C sucrose.

In this flux analysis study, comparisons of flux values in zygotic and SHaM embryos were based on relative flux values with respect to 100 moles of sucrose intake rate, representative of the carbon regulatory aspect of the metabolic pathway. For both embryo types, the total flow of carbon into biomass remained constant, and the flux maps indicated that the TCA cycle in mitochondrial, plastidic and cytosolic glycolysis, and lipid synthesis are significantly higher in zygotic than in SHaM embryos, as shown in Figure 4. However, the flux distribution of protein biosynthesis is similar between zygotic and SHaM embryos, except for the glutamine synthesis. The TCA cycle is commonly known for energy metabolism and ATP synthesis in plant system. In recent labeling studies, measured embryo cultures showed that the cycle is responsible, in terms of flux distribution toward carbon skeletons, for meeting biosynthesis requirements, and is dictated by the metabolic and physiological demands of soybean embryos (Allen *et al.*, 2009; Allen and Young, 2013; Iyer *et al.*, 2008; Sriram *et al.*, 2004), corn (Alonso *et al.*, 2010; Alonso *et al.*, 2007), Arabidopsis (Lonien and Schwender, 2009) . In our soybean studies, the zygotic embryos uptake more glutamine as a form of reduced nitrogen source, which resulting in the higher rates of transaminase reaction, and therefore act as carbon skeletons for 2-oxoglutrate, malate, and OAA. Hence, leveraging the respiratory CO_2 released in the system for zygotic embryos helps result in the increased metabolic activity of TCA cycle.

Furthermore, the high rate of glutamine uptake with a constant carbon source in zygotic embryos could possibly influence anaplerotic reactions. The anaplerotic reactions catalyzed by cytosolic ppc, mep, and malate (*MalT1* and *MalT2*)

transporters for shunting malate between cytosol, plastid, and mitochondrion were investigated. The flux through the anaplerotic reaction of pp^c was higher in zygotic embryos than in SHaM, as shown in Figure 4. The fluxes through plastidic and cytosolic malate were observed to be statistically similar between both embryos, indicating that the plastidic pyruvate flux must come from the glycolysis of the same compartment for both embryos. The low protein accumulation in SHaM embryos did not reflect in the low value of malic enzymes, which may indicate that the mechanism of protein synthesis in SHaM embryos does not depend on the malate enzyme in the reaction from plastidic malate to pyruvate. This, tied with the result that we observed from our in-depth study of SHaM embryos, reveals that the protein synthesis is strongly correlated with the glutamine level in the media (data not shown). A similar result was observed for the soy zygotic embryos studied by ^{13}C MFA (Allen and Young, 2013). Moreover, for both embryos, malic enzymes provide a small fraction of the pyruvate entering the TCA cycle, which is consistent with previous soy zygotic embryos studies (Allen *et al.*, 2009; Allen and Young, 2013; Iyer *et al.*, 2008). Therefore, the fluxes through the TCA cycle and those enzymes could be potential metabolic engineering targets for improving protein biosynthesis within the soybean developing embryos. One of the important branch points for understanding the carbon partitioning in protein and oil is flux through pyruvate and to acetyl-CoA in the plastid. Carbon in this pathway can be distributed to either oil or amino acid synthesis as shown in Figure 4; however, the flux into acetyl-CoA for oil synthesis is significantly higher for both soybean embryos, and especially higher for the zygotic embryos. The high expense of oil biosynthesis for zygotic embryos solidly depends

on the plastidic pyruvate dehydrogenase, which demand more ATP from the system. As a result, the glycolysis is sufficient to supply the ATP required for biomass production.

Acknowledgements

The authors would like to thank Christina M. Kostow for mature somatic embryos prior to labeling experiment. We would like to acknowledge Dr. D. Bruce Fulton (Department of Biochemistry, Biophysics and Molecular Biology, Iowa State University) for help with the NMR experiments and Dr. Jong Moon Yoon for training on NMR preparation samples and running simulation. This work was supported by DuPont Pioneer Hi-Bred.

References

Agrawal GK, Hajduch M, Graham K, Thelen JJ. 2008. In-depth investigation of the soybean seed-filling proteome and comparison with a parallel study of rapeseed. *Plant Physiology* 148, 504-518.

Allen DK, Libourel IGL, Shachar-Hill Y. 2009. Metabolic flux analysis in plants: coping with complexity. *Plant Cell and Environment* 32, 1241-1257.

Allen DK, Ohlrogge JB, Shachar-Hill Y. 2009. The role of light in soybean seed filling metabolism. *Plant Journal* 58, 220-234.

Allen DK, Shachar-Hill Y, Ohlrogge JB. 2007. Compartment-specific labeling information in C-13 metabolic flux analysis of plants. *Phytochemistry* 68, 2197-2210.

Allen DK, Young JD. 2013. Carbon and Nitrogen Provisions Alter the Metabolic Flux in Developing Soybean Embryos. *Plant Physiology* 161, 1458-1475.

Allen DK, Young JD. 2013. Carbon and Nitrogen Provisions Alter the Metabolic Flux in Developing Soybean Embryos. *Plant Physiology*.

Allen SM, Butler KH, Carlson TJ, Hitz WD, Stoop JM. 2008. Plastidic phosphoglucomutase genes. Vol. US 20080066204A1: EI du Pont de Nemours and Company.

Alonso AP, Dale VL, Shachar-Hill Y. 2010. Understanding fatty acid synthesis in developing maize embryos using metabolic flux analysis. *Metabolic Engineering* 12, 488-497.

Alonso AP, Goffman FD, Ohlrogge JB, Shachar-Hill Y. 2007. Carbon conversion efficiency and central metabolic fluxes in developing sunflower (Helianthus annuus L.) embryos. *Plant Journal* 52, 296-308.

Alonso AP, Piasecki RJ, Wang Y, LaClair RW, Shachar-Hill Y. 2010. Quantifying the Labeling and the Levels of Plant Cell Wall Precursors Using Ion Chromatography Tandem Mass Spectrometry. *Plant Physiology* 153, 915-924.

Alonso AP, Raymond P, Hernould M, et al. 2007. A metabolic flux analysis to study the role of sucrose synthase in the regulation of the carbon partitioning in central metabolism in maize root tips. *Metabolic Engineering* 9, 419-432.

Alonso AP, Raymond P, Rolin D, Dieuaide-Noubhani M. 2007. Substrate cycles in the central metabolism of maize root tips under hypoxia. *Phytochemistry* 68, 2222-2231.

Alonso AP, Val DL, Shachar-Hill Y. 2011. Central metabolic fluxes in the endosperm of developing maize seeds and their implications for metabolic engineering. *Metabolic Engineering* 13, 96-107.

Bailey MA, Boerma HR, Parrott WA. 1993. Genotype-Specific Optimization of Plant-Regeneration from Somatic Embryos of Soybean. *Plant Science* 93, 117-120.

Bodenhausen G, Ruben DJ. 1980. Natural abundance nitrogen-15 NMR by enhanced heteronuclear spectroscopy. *Chemical Physics Letters* 69, 185-189.

Boghigian BA, Seth G, Kiss R, Pfeifer BA. 2010. Metabolic flux analysis and pharmaceutical production. *Metabolic Engineering* 12, 81-95.

Borisjuk L, Nguyen TH, Neuberger T, et al. 2005. Gradients of lipid storage, photosynthesis and plastid differentiation in developing soybean seeds. *New Phytologist* 167, 761-776.

Bowsher CG, Tobin AK. 2001. Compartmentation of metabolism within mitochondria and plastids. *Journal of Experimental Botany* 52, 513-527.

Boyle NR, Morgan JA. 2009. Flux balance analysis of primary metabolism in Chlamydomonas reinhardtii. *BMC Systems Biology* 3, (7 January 2009).

Bradford MM. 1976. Rapid and Sensitive Method for Quantitation of Microgram Quantities of Protein Utilizing Principle of Protein-Dye Binding. *Analytical Biochemistry* 72, 248-254.

Brown LR. 1984. Differential Scaling along Omega-1 in COSY Experiments *Journal of Magnetic Resonance* 57, 513-518.

Carrari F, Urbanczyk-Wochniak E, Willmitzer L, Fernie AR. 2003. Engineering central metabolism in crop species: learning the system. *Metabolic Engineering* 5, 191-200.

Chanprame S, Kuo TM, Widholm JM. 1998. Soluble Carbohydrate Content of Soybean [Glycine max (L.) Merr.] Somatic and Zygotic Embryos during Development. *In Vitro Cellular & Developmental Biology. Plant* 34, 64-68.

Clemente TE, Cahoon EB. 2009. Soybean Oil: Genetic Approaches for Modification of Functionality and Total Content. *Plant Physiology* 151, 1030-1040.

Cohen SA. 2000. Amino Acid Analysis Using Precolumn Derivatization with 6-Aminoquinolyl-N-Hydroxysuccinimidyl Carbamate. *Methods in Molecular Biology: Amino Acid Analysis Protocols* 159, 39-47.

Dietz K. 1987. Control function of hexosemonophosphate isomerase and phosphoglucomutase in starch synthesis in leaves. *In J Biggins, ed, Proceedings of the 7th International Congress on Photosynthesis* 3, 329-332.

Dieuaide-Noubhani M, Alonso AP, Rolin D, Eisenreich W, Raymond P. 2007. *Metabolic flux analysis: recent advances in carbon metabolism in plants*.

Durot M, Bourguignon PY, Schachter V. 2009. Genome-scale models of bacterial metabolism: reconstruction and applications. *FEMS Microbiology Reviews* 33, 164-190.

Fait A, Fromm H, Walter D, Galili G, Fernie AR. 2008. Highway or byway: the metabolic role of the GABA shunt in plants. *Trends in Plant Science* 13, 14-19.

Feist AM, Herrgard MJ, Thiele I, Reed JL, Palsson BO. 2009. Reconstruction of biochemical networks in microorganisms. *Nature Reviews Microbiology* 7, 129-143.

Finer JJ, McMullen MD. 1991. Transformation of Soybean Via Particle Bombardment of Embryogenic Suspension Culture Tissue. *In Vitro Cellular and Developmental Biology Plant* 27P, 175-182.

Gifford RM, Thorne JH. 1985. Sucrose Concentration at the Apoplastic Interface Between Seed Coat and Cotyledons of Developing Soybean Seeds. *Plant Physiology* 77, 863-868.

Golombek S, Rolletschek H, Wobus U, Weber H. 2001. Control of storage protein accumulation during legume seed development. *Journal of Plant Physiology* 158, 457-464.

Gonzalez DO, Vodkin LO. 2007. Specific elements of the glyoxylate pathway play a significant role in the functional transition of the soybean cotyledon during seedling development. *Bmc Genomics* 8.

Haga KI, Sodek L. 1987. Utilization of Nitrogen-Sources by Immature Soybean Cotyledons in Culture. *Annals of Botany* 59, 597-601.

Harris RK. 1983. Nuclear Magnetic Resonance Spectroscopy: A Physiochemical View. *Pitman Books, London*.

Hartwig EE, Kilen TC. 1991. Yield and Composition of Soybean Seed from Parents with Different Protein, Similar Yield. *Crop Science* 31, 290-292.

Hattenbach A, Heineke D. 1999. On the role of chloroplastic phosphoglucomutase in the regulation of starch turnover. *Planta* 207, 527-532.

Hay J, Schwender J. 2011. Computational analysis of storage synthesis in developing Brassica napus L. (oilseed rape) embryos: flux variability analysis in relation to (13)C metabolic flux analysis. *Plant Journal* 67, 513-525.

Hay J, Schwender J. 2011. Metabolic network reconstruction and flux variability analysis of storage synthesis in developing oilseed rape (Brassica napus L.) embryos. *Plant Journal* 67, 526-541.

Hayati R, Egli DB, CraftsBrandner SJ. 1996. Independence of nitrogen supply and seed growth in soybean: Studies using an in vitro culture system. *Journal of Experimental Botany* 47, 33-40.

He Y, Young T, Clark K, Kleppinger-Sparace K, Bridges W, Sparace S. 2011. Developmental profile of storage reserve accumulation in soybean somatic embryos. *In Vitro Cellular & Developmental Biology - Plant* 47, 725-733.

Hernandez-Sebastia C, Marsolais F, Saravitz C, Israel D, Dewey RE, Huber SC. 2005. Free amino acid profiles suggest a possible role for asparagine in the control of storage-product accumulation in developing seeds of low- and high-protein soybean lines. *Journal of Experimental Botany* 56, 1951-1963.

Hsu FC, Bennett AB, Spanswick RM. 1984. Concentrations of Sucrose and Nitrogenous Compounds in the Apoplast of Developing Soybean Seed Coats and Embryos *Plant Physiology* 75, 181-186.

Hymowitz T, Collins FI, Walker WM, Panczner J. 1972. Relationship between content of oil, protein, and sugar in soybean seed. *Agronomy Journal* 64, 613-616.

Iyer VV, Sriram G, Fulton DB, Zhou R, Westgate ME, Shanks JV. 2008. Metabolic flux maps comparing the effect of temperature on protein and oil biosynthesis in developing soybean cotyledons. *Plant Cell and Environment* 31, 506-517.

Iyer VV, Sriram G, Shanks JV. 2007. Metabolic flux maps of central carbon metabolism in plant systems - Isotope labeling analysis. *Concepts in Plant Metabolomics*, 125-144.

Johnson BA, Blevins RA. 1994. NMR View - A Computure-Program for the Visualization and Analysis of NMR data. *Journal of Biomolecular Nmr* 4, 603-614.

Joshi V, Laubengayer KM, Schauer N, Fernie AR, Jander G. 2006. Two Arabidopsis threonine aldolases are nonredundant and compete with threonine deaminase for a common substrate pool. *Plant Cell* 18, 3564-3575.

Junker BH, Lonien J, Heady LE, Rogers A, Schwender J. 2007. Parallel determination of enzyme activities and in vivo fluxes in Brassica napus embryos grown on organic or inorganic nitrogen source. *Phytochemistry* 68, 2232-2242.

Kell DB. 2004. Metabolomics and systems biology: making sense of the soup. *Current Opinion in Microbiology* 7, 296-307.

Kinney AJ. 1996. Development of genetically engineered soybean oils for food applications. *Journal of Food Lipids* 3, 273-292.

Koch K. 2004. Sucrose metabolism: regulatory mechanisms and pivotal roles in sugar sensing and plant development. *Current Opinion in Plant Biology* 7, 235-246.

Koch KE. 1996. Carbohydrate-modulated gene expression in plants. *Annual Review of Plant Physiology and Plant Molecular Biology* 47, 509-540.

Krivdin LB, Kalabin GA. 1989. Structural Applications of One-Bond Carbon-Carbon Spin-Spin Coupling-Constants *Progress in Nuclear Magnetic Resonance Spectroscopy* 21, 293-448.

Kruger NJ, Masakapalli SK, Ratcliffe RG. 2012. Strategies for investigating the plant metabolic network with steady-state metabolic flux analysis: lessons from an Arabidopsis cell culture and other systems. *Journal of Experimental Botany* 63, 2309-2323.

Kruger NJ, Ratcliffe RG. 2009. Insights into plant metabolic networks from steady-state metabolic flux analysis. *Biochimie* 91, 697-702.

Kruger NJ, Troncoso-Ponce MA, Ratcliffe RG. 2008. 1H NMR metabolite fingerprinting and metabolomic analysis of perchloric acid extracts from plant tissues. *Nat. Protocols* 3, 1001-1012.

Kruger NJ, von Schaewen A. 2003. The oxidative pentose phosphate pathway: structure and organisation. *Current Opinion in Plant Biology* 6, 236-246.

Last RL, Jones AD, Shachar-Hill Y. 2007. Towards the plant metabolome and beyond. *Nature Reviews Molecular Cell Biology* 8, 167-174.

Li ZS, Moon BP, Xing AQ, Liu ZB, McCardell RP, Damude HG, Falco SC. 2010. Stacking Multiple Transgenes at a Selected Genomic Site via Repeated Recombinase-Mediated DNA Cassette Exchanges. *Plant Physiology* 154, 622-631.

Libourel IGL, Shachar-Hill Y. 2008. Metabolic flux analysis in plants: from intelligent design to rational engineering. *Annual Review of Plant Biology* 59, 625-650.

Litterer LA, Plaisance KL, Schnurr JA, Storey KK, Jung HJG, Gronwald JW, Somers DA. 2006. Biosynthesis of UDP-glucuronic acid in developing soybean embryos: possible role of UDP-sugar pyrophosphorylase. *Physiologia Plantarum* 128, 200-211.

Lonien J, Schwender J. 2009. Analysis of Metabolic Flux Phenotypes for Two Arabidopsis Mutants with Severe Impairment in Seed Storage Lipid Synthesis. *Plant Physiology* 151, 1617-1634.

Lu C, Napier JA, Clemente TE, Cahoon EB. 2011. New frontiers in oilseed biotechnology: meeting the global demand for vegetable oils for food, feed, biofuel, and industrial applications. *Current Opinion in Biotechnology* 22, 252-259.

Manjunath S, Lee CHK, VanWinkle P, Bailey-Serres J. 1998. Molecular and biochemical characterization of cytosolic phosphoglucomutase in maize - Expression during development and in response to oxygen deprivation. *Plant Physiology* 117, 997-1006.

Masakapalli SK, Lay PL, Huddleston JE, Pollock NL, Kruger NJ, Ratcliffe RG. 2009. Subcellular flux analysis of central metabolism in a heterotrophic Arabidopsis thaliana cell suspension using steady-state stable isotope labeling. *Plant Physiology*, pp.109.151316.

Masakapalli SK, Le Lay P, Huddleston JE, Pollock NL, Kruger NJ, Ratcliffe RG. 2010. Subcellular Flux Analysis of Central Metabolism in a Heterotrophic Arabidopsis Cell Suspension Using Steady-State Stable Isotope Labeling. *Plant Physiology* 152, 602-619.

Massou S, Nicolas C, Letisse F, Portais JC. 2007. NMR-based fluxomics: Quantitative 2D NMR methods for isotopomers analysis. *Phytochemistry* 68, 2330-2340.

Medini D, Serruto D, Parkhill J, Relman DA, Donati C, Moxon R, Falkow S, Rappuoli R. 2008. Microbiology in the post-genomic era. *Nature Reviews Microbiology* 6, 419-430.

Meyer K, Hitz WD, Yadav NS, Damude HG. 2012. DGAT genes from Yarrowia lipolytica for increased seed storage lipid production and altered fatty acid profiles in soybean. E I du Pont de Nemours and Company. US Patent No US 8143473.

Meyer K, Stecca KL, Ewell-Hicks K, Allen SM, Everard JD. 2012. Oil and Protein Accumulation in Developing Seeds Is Influenced by the Expression of a Cytosolic Pyrophosphatase in Arabidopsis. *Plant Physiology* 159, 1221-1234.

Miflin BJ, Lea PJ. 1977. Amino- Acid Metabolism *Briggs, Winslow R. (Ed.). Annual Review of Plant Physiology, Vol. 28. X+615p. Illus. Annual Reviews Inc.: Palo Alto, Calif., USA. Isbn 0-8243-0628-7*, 299-329.

Miflin BJ, Lea PJ. 1977. Amino Acid Metabolism. *Annual Review of Plant Physiology* 28, 299-329.

Miranda M, Borisjuk L, Tewes A, Heim U, Sauer N, Wobus U, Weber H. 2001. Amino acid permeases in developing seeds of Vicia faba L.: expression precedes storage protein synthesis and is regulated by amino acid supply. *Plant Journal* 28, 61-71.

Morgan JA, Barney CS, Penn AH, Shanks JV. 2000. Effects of buffered media upon growth and alkaloid production of Catharanthus roseus hairy roots. *Applied Microbiology and Biotechnology* 53, 262-265.

Nargund S, Sriram G. 2013. Designer labels for plant metabolism: statistical design of isotope labeling experiments for improved quantification of flux in complex plant metabolic networks. *Molecular Biosystems* 9, 99-112.

Nielsen J, Oliver S. 2005. The next wave in metabolome analysis. *Trends in Biotechnology* 23, 544-546.

Nishizawa K, Ishimoto M. 2009. Maturation of somatic embryos as a model for soybean seed development. *Plant Biotechnology* 26, 543-550.

Nishizawa K, Takagi K, Teraishi M, Kita A, Ishimoto M. 2010. Application of somatic embryos to rapid and reliable analysis of soybean seed components by RNA interference-mediated gene silencing. *Plant Biotechnology* 27, 409-420.

Nishizawaa K, Ishimoto M. 2009. Maturation of somatic embryos as a model for soybean seed development. *Plant Biotechnology* 26, 543-550.

Norton G, Harris JF. 1975. Compositional Changes in Developing Rape Seed (Brassica-Napus L) *Planta* 123, 163-174.

O'Grady J, Schwender J, Shachar-Hill Y, Morgan JA. 2012. Metabolic cartography: experimental quantification of metabolic fluxes from isotopic labelling studies. *Journal of Experimental Botany* 63, 2293-2308.

Obendorf RL, Rytko GT, Byrne MC. 1982. Soya Bean Seed Growth and Maturation by In Vitro Pod Culture. *Ann Bot* 51, 217-227.

Obendorf RL, Rytko GT, Byrne MC. 1983. Soya Bean Seed Growth and Maturation by In Vitro Pod Culture. *Ann Bot* 51, 217-227.

Pandurangan S, Pajak A, Molnar SJ, et al. 2012. Relationship between asparagine metabolism and protein concentration in soybean seed. *Journal of Experimental Botany* 63, 3173-3184.

Paula Alonso A, Dale VL, Shachar-Hill Y. 2010. Understanding fatty acid synthesis in developing maize embryos using metabolic flux analysis. *Metabolic Engineering* 12, 488-497.

Periappuram C, Steinhauer L, Barton DL, Taylor DC, Chatson B, Zou J. 2000. The plastidic phosphoglucomutase from Arabidopsis. A reversible enzyme reaction with an important role in metabolic control. *Plant Physiology* 123, 1197-1197.

Pipolo AE, Sinclair TR, Camara GMS. 2004. Protein and oil concentration of soybean seed cultured in vitro using nutrient solutions of differing glutamine concentration. *Annals of Applied Biology* 144, 223-227.

Ratcliffe RG, Shachar-Hill Y. 2001. Probing plant metabolism with NMR. *Annual Review of Plant Physiology and Plant Molecular Biology* 52, 499-526.

Ratcliffe RG, Shachar-Hill Y. 2006. Measuring multiple fluxes through plant metabolic networks. *Plant Journal* 45, 490-511.

Rolletschek H, Radchuk R, Klukas C, Schreiber F, Wobus U, Borisjuk L. 2005. Evidence of a key role for photosynthetic oxygen release in oil storage in developing soybean seeds. *New Phytologist* 167, 777-786.

Rontein D, Dieuaide-Noubhani M, Dufourc EJ, Raymond P, Rolin D. 2002. The metabolic architecture of plant cells - Stability of central metabolism and flexibility of anabolic pathways during the growth cycle of tomato cells. *Journal of Biological Chemistry* 277, 43948-43960.

Ruuska SA, Schwender J, Ohlrogge JB. 2004. The capacity of green oilseeds to utilize photosynthesis to drive biosynthetic processes. *Plant Physiology* 136, 2700-2709.

Salon C, Munier-Jolain NG, Duc G, Voisin AS, Grandgirard D, Larmure A, Emery RJN, Ney B. 2001. Grain legume seed filling in relation to nitrogen acquisition: A review and prospects with particular reference to pea. *Agronomie* 21, 539-552.

Saravitz CH, Raper CD. 1995. Responses to Sucrose and Glutamine by Soybean Embryos Growth *In-Vitro*. *Physiologia Plantarum* 93, 799-805.

Schmidt MA, Tucker DM, Cahoon EB, Parrott WA. 2005. Towards normalization of soybean somatic embryo maturation. *Plant Cell Reports* 24, 383-391.

Schmutz J, Cannon SB, Schlueter J, et al. 2010. Genome sequence of the palaeopolyploid soybean. *Nature* 463, 178-183.

Schnarrenberger C, Oeser A, Tolbert NE. 1973. Two isoenzymes each of glucose-6-phosphate dehydrogenase and 6-phosphogluconate dehydrogenase in spinach leaves. *Archives of Biochemistry and Biophysics* 154, 438-448.

Schwender J. 2008. Metabolic flux analysis as a tool in metabolic engineering of plants. *Current Opinion in Biotechnology* 19, 131-137.

Schwender J, Goffman F, Ohlrogge JB, Shachar-Hill Y. 2004. Rubisco without the Calvin cycle improves the carbon efficiency of developing green seeds. *Nature* 432, 779-782.

Schwender J, Ohlrogge JB. 2002. Probing in vivo metabolism by stable isotope labeling of storage lipids and proteins in developing Brassica napus embryos. *Plant Physiology* 130, 347-361.

Schwender J, Ohlrogge JB, Shachar-Hill Y. 2003. A flux model of glycolysis and the oxidative pentosephosphate pathway in developing Brassica napus embryos. *Journal of Biological Chemistry* 278, 29442-29453.

Schwender J, Shachar-Hill Y, Ohlrogge JB. 2006. Mitochondrial metabolism in developing embryos of Brassica napus. *Journal of Biological Chemistry* 281, 34040-34047.

Shachar-Hill Y. 2002. Nuclear magnetic resonance and plant metabolic engineering. *Metabolic Engineering* 4, 90-97.

Sinclair TR, de Wit CT. 1975. Photosynthate and Nitrogen Requirements for Seed Production by Various Crops. *Science* 189, 565-567.

Smeekens S. 2000. Sugar-induced signal transduction in plants. *Annual Review of Plant Physiology and Plant Molecular Biology* 51, 49-81.

Smith AJ, Rinne RW, Seif RD. 1989. Phosphoenolpyruvate Carboxylase and Pyruvate Kinase Involvement in Protein and Oil Biosynthesis during Soybean Seed Development. *Crop Sci.* 29, 349-353.

Sriram G, Fulton DB, Iyer VV, Peterson JM, Zhou R, Westgate ME, Spalding MH, Shanks JV. 2004. Quantification of compartmented metabolic fluxes in developing soybean embryos by employing biosynthetically directed fractional ^{13}C labeling, two-dimensional [^{13}C, ^{1}H] nuclear magnetic resonance, and comprehensive isotopomer balancing. *Plant Physiology* 136, 3043-3057.

Sriram G, Fulton DB, Iyer VV, Peterson JM, Zhou RL, Westgate ME, Spalding MH, Shanks JV. 2004. Quantification of compartmented metabolic fluxes in developing soybean embryos by employing Biosynthetic ally directed fractional C-13 labeling,

C-13, H-1 two-dimensional nuclear magnetic resonance, and comprehensive isotopomer balancing. *Plant Physiology* 136, 3043-3057.

Sriram G, Fulton DB, Shanks JV. 2007. Flux quantification in central carbon metabolism of Catharanthus roseus hairy roots by C-13 labeling and comprehensive bondomer balancing. *Phytochemistry* 68, 2243-2257.

Sriram G, Iyer VV, Fulton DB, Shanks JV. 2007. Identification of hexose hydrolysis products in metabolic flux analytes: A case study of levulinic acid in plant protein hydrolysate. *Metabolic Engineering* 9, 442-451.

Stepansky A, Leustek T. 2006. Histidine biosynthesis in plants. *Amino Acids* 30, 127-142.

Stephanopoulos G. 1999. Metabolic Fluxes and Metabolic Engineering. *Metabolic Engineering* 1, 1-11.

Stephanopoulos G. 2002. Metabolic engineering: Perspective of a chemical engineer. *Aiche Journal* 48, 920-926.

Stephanopoulos G, Vallino JJ. 1991. Network Rigidity and Metabolic Engineering in Metabolite Overproduction. *Science* 252, 1675-1681.

Stephanopoulos GN, Aristidou AA, Nielsen J. 1998. Metabolic Flux Analysis. *Metabolic Engineering*. San Diego: Academic Press, 309-351.

Stombaugh SK, Jung HG, Orf JH, Somers DA. 2000. Genotypic and environmental variation in soybean seed cell wall polysaccharides. *Crop Science* 40, 408-412.

Stombaugh SK, Orf JH, Jung HG, Somers DA. 2003. Relationships between soybean seed cell wall polysaccharides, yield, and seed traits. *Crop Science* 43, 571-576.

Sweetlove LJ, Fell D, Fernie AR. 2008. Getting to grips with the plant metabolic network. *Biochem J* 409, 27-41.

Szyperski T. 1995. Biosynthetically directed fractional ^{13}C-labeling of proteinogenic amino acids. An efficient analytical tool to investigate intermediary metabolism. *European Journal of Biochemistry* 232, 433-448.

Szyperski T. 1995. Biosynthetically Directed Fractional C-13-Labeling of Proteinogenic Amino-Acids - An Efficient Analytical Tool to Investigate Intermediary Metabolism. *European Journal of Biochemistry* 232, 433-448.

Thompson JF, Madison JT, Muenster AME. 1977. *In Vitro* Culture of Immature Cotyledons of Soybean (*Glycine-Max-L-Merr*) *Annals of Botany* 41, 29-39.

Troufflard S, Roscher A, Thomasset B, Barbotin JN, Rawsthorne S, Portais JC. 2007. In vivo C-13 NMR determines metabolic fluxes and steady state in, linseed embryos. *Phytochemistry* 68, 2341-2350.

van der Werf MJ, Overkamp KM, Muilwijk B, Coulier L, Hankemeier T. 2007. Microbial metabolomics: Toward a platform with full metabolome coverage. *Analytical Biochemistry* 370, 17-25.

van Winden W, Schipper D, Verheijen P, Heijnen J. 2001. Innovations in generation and analysis of 2D C-13, H-1 COSYNMR spectra for metabolic flux analysis purposes. *Metabolic Engineering* 3, 322-343.

van Winden WA, Heijnen JJ, Verheijen PJT, Grievink J. 2001. A priori analysis of metabolic flux identifiability from C-13-labeling data. *Biotechnology and Bioengineering* 74, 505-516.

Varma A, Palsson BO. 1994. Stoichiometric Flux Balance Models Quantitatively Predict Growth and Metabolic By-Product Secretion in Wild-Type *Escherichia-Coli W3110*. *Applied and Environmental Microbiology* 60, 3724-3731.

Wakao S, Andre C, Benning C. 2008. Functional analyses of cytosolic glucose-6-phosphate dehydrogenases and their contribution to seed oil accumulation in Arabidopsis. *Plant Physiology* 146, 277-288.

Walker DR, Parrott WA. 2001. Effect of polyethylene glycol and sugar alcohols on soybean somatic embryo germination and conversion. *Plant Cell Tissue and Organ Culture* 64, 55-62.

Weigelt K, Kuster H, Radchuk R, Muller M, Weichert H, Fait A, Fernie AR, Saalbach I, Weber H. 2008. Increasing amino acid supply in pea embryos reveals specific interactions of N and C metabolism, and highlights the importance of mitochondrial metabolism. *Plant Journal* 55, 909-926.

Wiechert W. 2001. C-13 metabolic flux analysis. *Metabolic Engineering* 3, 195-206.

Wiechert W, Mollney M, Petersen S, de Graaf AA. 2001. A universal framework for C-13 metabolic flux analysis. *Metabolic Engineering* 3, 265-283.

Wilcox JR, Shibles RM. 2001. Interrelationships among seed quality attributes in soybean. *Crop Science* 41, 11-14.

Williams TCR, Miguet L, Masakapalli SK, Kruger NJ, Sweetlove LJ, Ratcliffe RG. 2008. Metabolic network fluxes in heterotrophic arabidopsis cells: stability of the flux distribution under different oxygenation conditions. *Plant Physiology* 148, 704-718.

Willker W, Flogel U, Leibfritz D. 1997. Ultra-high-resolved HSQC spectra of multiple-C-13-labeled biofluids. *Journal of Magnetic Resonance* 125, 216-219.

Wilson LA. 1995. Soy foods. In D. R. Erickson (ed.) Practial handbook of soybean processing and utilization. AOCS Press, Champaign, IL and United Soybean Board, St. Louis, MO., 428-459.

Wuthrich K. 1976. NMR in Biological Research: Peptides and Proteins. *North Holland, Amsterdam*.

Figures caption

Figure 1: Comparison of nonlabeled and labeled sucrose cultured of soybean zygotic and SHaM embryos on biomass composition of total oil, protein, soluble sugar, and residual biomass. Significant differences between the labeled culture of zygotic and SHaM embryos are indicated according to Student's mean comparison test (* $P < 0.05$).

Figure 2: Correlation between the relative multiplet 2D [^{13}C, ^{1}H] HSQC NMR intensities of individual amino acids in hydrolyzated protein from (A) zygotic and SHaM embryos (*cv. Jack*). Gray shading indicates the 95% confidence and prediction intervals of the relative intensities.

Figure 3: Comparison of corresponding multiplet intensities of glucosyl units (HMF1 and LVA 3, LVA 6) hydrolysate of protein (cytosol compartment) and starch (plastidic compartment) of zygotic and SHaM embryos *cv. Jack*. Gray shading indicates the 95% confidence and prediction intervals of the relative intensities.

Figure 4: Comparison of metabolic flux map of SHaM embryos to zygotic embryos (*cv. Jack)* with relative units (normalized to 100 mole of sucrose uptake)*;* the map representing central carbon metabolism of the soybean model constructed for the earlier work of developing soybean cotyledon (Sriram et al., 2004). Significant differences between the two embryos are displayed according to Student's t mean comparison test ($P < 0.10$). The arrows thickness represent the differences in fluxes

and colored of the arrow indicate the flux of low flux (red) and high flux (blue) as flux values were compared from SHaM to zygotic embryos.

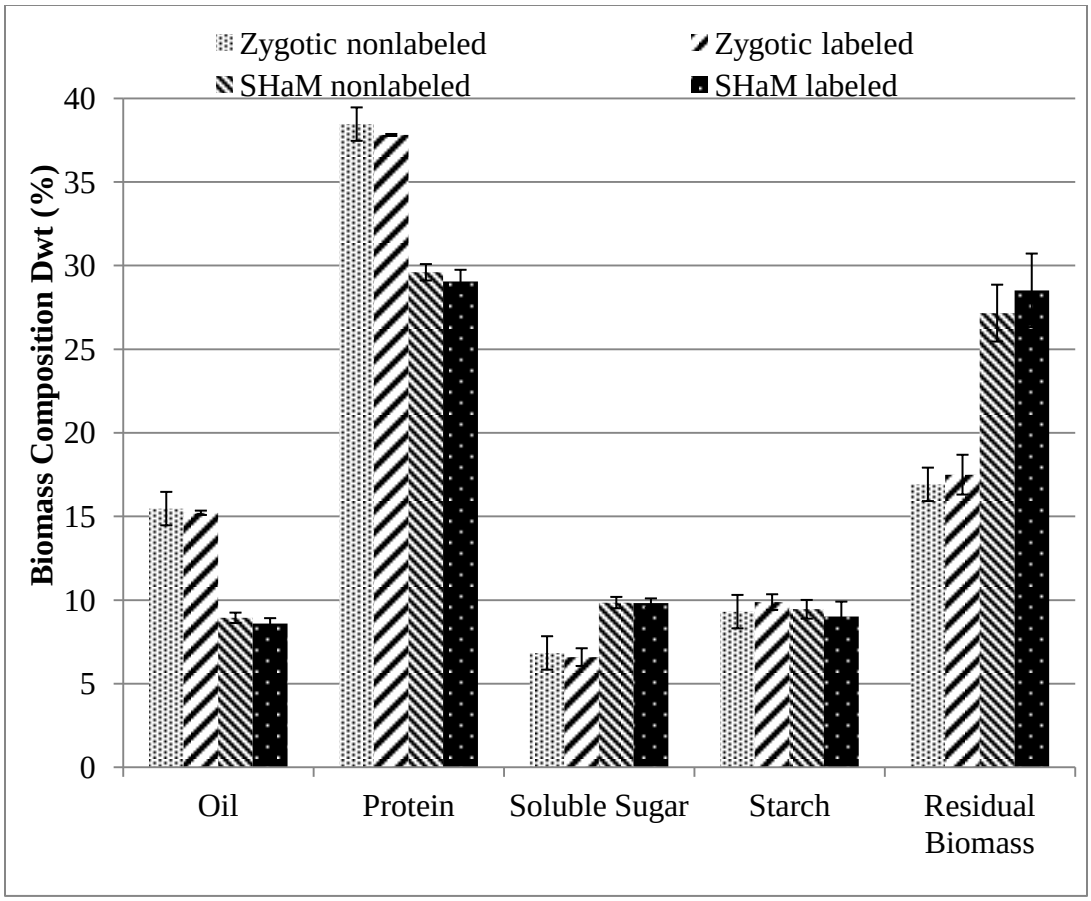

Zygotic nonlabeled
Zygotic labeled
SHaM nonlabeled
SHaM labeled
Biomass Composition Dwt (%)
40
35
30
25
20
15
10
5
0
Oil
Protein
Soluble Sugar
Starch
Residual Biomass

Figure 1

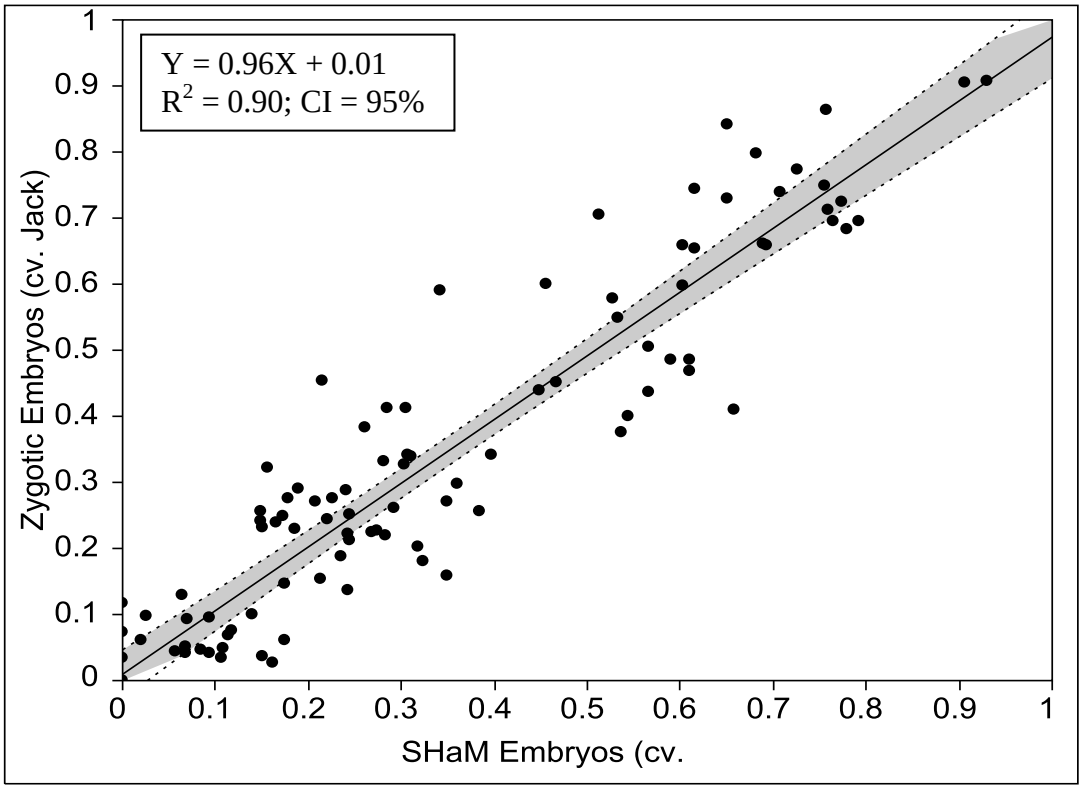


Figure 2

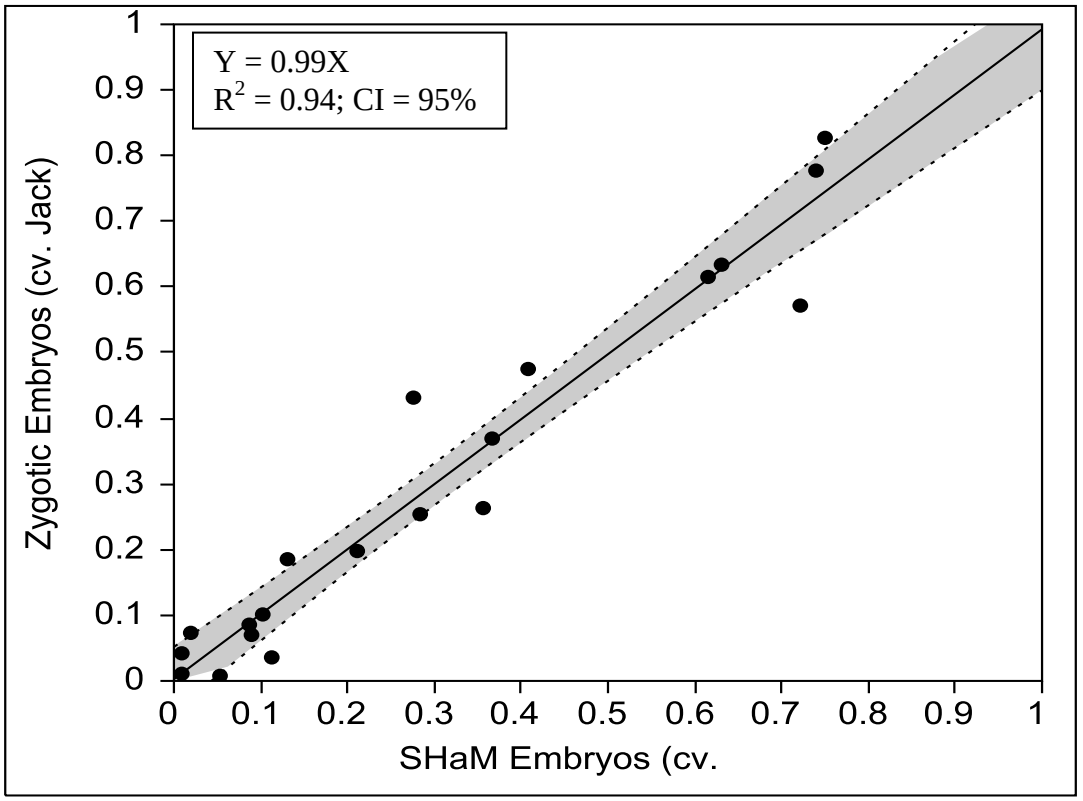


Figure 3

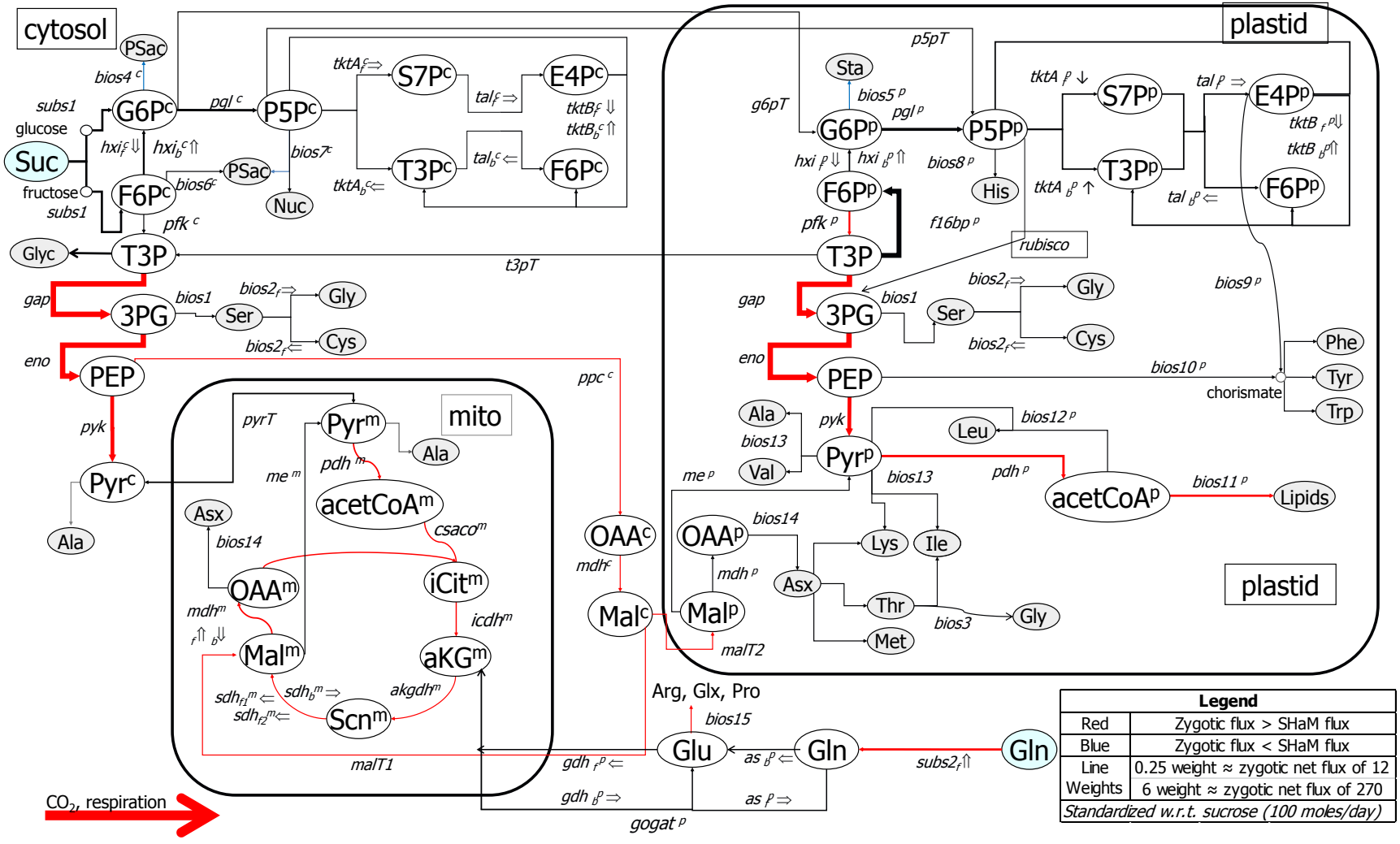

cytosol
plastid
mito
PSac
bios4 c
subs1
glucose
Suc
fructose
subs1
G6Pc
hxi f c ⇓
hxi b c ⇑
F6Pc
pgl c
P5Pc
bios6 c
PSac
Nuc
bios7 c
tktA f c ⇒
S7Pc
tktA b c ⇐
T3Pc
tal f c ⇒
tal b c ⇐
E4Pc
F6Pc
tktB f c ⇓
tktB b c ⇑
pfk c
Glyc
T3P
t3pT
gap
3PG
bios1
Ser
bios2 f ⇒
Gly
bios2 f ⇐
Cys
eno
PEP
pyk
Pyrc
Ala
ppc c
pyrT
Pyrm
Ala
me m
pdh m
acetCoAm
csaco m
Asx
bios14
OAAm
iCitm
mdhm
icdhm
f ⇑ b ⇓
Malm
aKGm
sdh f1 m ⇐
sdh f2 m ⇐
sdh b m ⇒
akgdh m
Scnm
malT1
CO2, respiration
OAAc
mdhc
Malc
p5pT
Sta
bios5 p
g6pT
G6Pp
pgl p
P5Pp
hxi f p ⇓
hxi b p ⇑
bios8 p
His
F6Pp
pfk p
f16bp p
T3P
rubisco
tktA f p ↓
S7Pp
T3Pp
tktA b p ↑
tal f p ⇒
E4Pp
F6Pp
tktB f p ⇓
tktB b p ⇑
tal b p ⇐
bios2 f ⇒
Gly
Ser
Cys
bios2 f ⇐
bios9 p
gap
3PG
bios1
eno
PEP
bios10 p
Phe
Tyr
Trp
chorismate
pyk
Ala
bios13
Val
Pyrp
me p
bios13
pdh p
Leu
bios12 p
acetCoAp
bios11 p
Lipids
OAAp
bios14
mdh p
Malp
malT2
Asx
Lys
Ile
Thr
Met
bios3
Gly
plastid
Arg, Glx, Pro
bios15
Glu
Gln
Gln
as b p ⇐
as f p ⇒
subs2 f ⇑
gdh f p ⇐
gdh b p ⇒
gogat p
Legend
Red
Zygotic flux > SHaM flux
Blue
Zygotic flux < SHaM flux
Line Weights
0.25 weight ≈ zygotic net flux of 12
6 weight ≈ zygotic net flux of 270
Standardized w.r.t. sucrose (100 moles/day)

Figure 4

Table 1: Normalizing growth rate based on dry weight and sucrose and glutamine uptake rates of zygotic and SHaM embryos (*cv. Jack*) cultured in labeled sucrose for 6 day with media exchange every three days. *Significantly different with Student't t mean comparison test $P<0.05$

Soybean *in vitro studies* cultivar	**Zygotic Embryos (*cv. Jack*)**	**SHaM Embryos (*cv. Jack)***
Relative growth rate (1/day)	0.52 ± 0.10	0.62 ± 0.03
Sucrose uptake (umol/g Dwt-day)	1360 ± 150	990 ± 54*
Glutamine uptake (umol/g Dwt-day)	750 ± 25	490 ± 20*

Table 2: Values of relative fluxes for the two embryo types analyzed in this study. Relative fluxes are normalized to 100 mole of sucrose uptake. Values show best fit of flux estimation with standard deviation as determined based on Monte Carlo simulation of 500 times. Boldface values indicate that the difference of flux between the zygotic and SHaM embryos is significant (P = 90%; Student t' test mean comparison with n = 3 for SHaM and n = 8 for zygotic).

Reaction	Reaction *Name*	Zygotic Embryos Standardized w.r.t sucrose (100 moles/day)		SHaM Embryos Standardized w.r.t sucrose (100 moles/day)	
		Net Flux	S.D.	Net Flux	S.D.
Glycolysis and oxPPP					
$G6P^c \leftrightarrow F6P^c$	hxi_f^c	-89	25	-88	12
reversibility %		0.58	0.22	0.59	0.27
$G6P^p \leftrightarrow F6P^p$	hxi_f^p	-66	11	-77	14
reversibility %		0.87	0.11	0.73	0.15
$G6P^c \rightarrow P5P^c + CO2$	pgl^c	114	31	119	22
$G6P^p \rightarrow P5P^p + CO2$	pgl^p	137	18	130	25
$P5P^c + P5P^c \leftrightarrow S7P^c + T3P^c$	$tktA_f^c$	9	5	6	3
reversibility %		0.87	0.14	0.81	0.08
$S7P^c + T3P^c \leftrightarrow F6P^c + E4P^c$	tal_f^c	9	5	6	3
reversibility %		0.99	0.01	0.88	0.10
$P5P^c + E4P^c \leftrightarrow F6P^c + T3P^c$	$tktB_f^c$	9	5	6	3
reversibility %		0.58	0.20	0.37	0.14
$P5P^p + P5P^p \leftrightarrow S7P^p + T3P^p$	$tktA_f^p$	54	10	70	6
reversibility %		0.97	0.04	0.56	0.43
$S7P^p + T3P^p \leftrightarrow F6P^p + E4P^p$	tal_f^p	54	10	70	6
reversibility %		0.83	0.05	0.55	0.23
$P5P^p + E4P^p \leftrightarrow F6P^p + T3P^p$	$tktB_f^p$	51	10	68	6
reversibility %		0.31	0.21	0.77	0.16
$F6P^c \rightarrow T3P^c + T3P^c$	pfk^c	13	5	26	10
$F6P^p \rightarrow T3P^p + T3P^p$	pfk^p	**86**	**5**	**39**	**13**
$T3P^p + T3P^p \rightarrow F6P^p$	$f16bp^p$	210	26	151	32
$T3P \rightarrow 3PG$	gap	**268**	**9**	**185**	**20**
$3PG \rightarrow PEP$	eno	**263**	**9**	**181**	**20**
$PEP \rightarrow Pyr$	pyk	**180**	**6**	**105**	**11**
$Pyr^p \rightarrow ACA^p + CO_2$	pdh^p	**111**	**4**	**60**	**7**
$Pyr^m \rightarrow ACA^m + CO_2$	pdh^m	58	3	64	4
$Pyr^c \rightarrow Pyr^m$	$pyrT$	68	8	71	3
$G6P^c \leftrightarrow G6P^p$	$g6pT_f$	48	22	21	12
reversibility %		0.43	0.08	0.50	0.25
$P5P^c \leftrightarrow P5P^p$	$p5pT_f$	22	20	65	19
reversibility %		0.09	0.17	0.81	0.19
$P5P^p + CO_2 \rightarrow 3PG + 3PG$	$rubisco$	9	5	7	3

Table 2: (Continued.)

		$T3P^c$	↔	$T3P^p$			*$t3pT_f$*	49	11	19	23
		reversibility %						0.93	0.03	0.31	0.28
Reaction							Reaction *Name*	Zygotic Embryos Standardized w.r.t sucrose (100 moles/day) Net Flux	S.D.	SHaM Embryos Standardized w.r.t sucrose (100 moles/day) Net Flux	S.D.
TCA cycle											
ACA^m	+	OAA^m	→	$ICit^m$			*$csaco^m$*	**56**	**8**	**33**	**5**
		$ICit^m$	→	aKG^m	+	CO_2	*$icdh^m$*	**56**	**8**	**33**	**5**
		aKG^m	→	Scn^m	+	CO_2	*$akgdh^m$*	**46**	**9**	**28**	**5**
		Scn^m	→	Mal^m			*$sdh_{f1}{}^m$*	**46**	**9**	**28**	**5**
		Scn^m	↔	Mal^m			*$sdh_{f2}{}^m$*	**46**	**9**	**92**	**12**
		reversibility %						0.85	0.04	0.39	0.03
		Mal^m	↔	OAA^m			*$mdh_f{}^m$*	**58**	**8**	**36**	**5**
		reversibility %						0.49	0.35	0.54	0.31
Anaplerotic reactions											
PEP^c	+	CO_2	→	OAA^c			*ppc^c*	**27**	**6**	**16**	**2**
		Mal^m	→	Pyr^m	+	CO_2	*me^m*	10	3	6	1
		Mal^p	→	Pyr^p	+	CO_2	*me^p*	6	1	4	1
Substrate entry											
		Suc^{ex}	→	$G6P^c$	+	$F6P^c$	*subs1*	100	0	100	0
		Gln^{ex}	↔	Gln^p			*$subs2_f$*	**65**	**1**	**50**	**1**
		reversibility %						0.99	0.01	0.90	0.01
Glutamate assimilation											
		Gln^p	→	Glu^p			*as_f*	58	2	67	10
		Glu^p	→	aKG^p			*$gdh_f{}^p$*	53	3	69	10
aKG^p	+	Gln^p	→	Glu^p	+	Glu^p	*$gogat^p$*	17	5	23	10
Malate shuttle											
		Mal^m	↔	Mal^c			*$malT1_f$*	**-8**	**1**	**-3**	**2**
		reversibility %						0.99	0.00	0.95	0.02
		Mal^c	→	Mal^p			*malT2*	**36**	**5**	**19**	**4**
		Mal^c	→	OAA^c			*mdh^c*	**27**	**6**	**16.10**	**1.50**
		Mal^p	→	OAA^p			*mdh^p*	13	1	14	1
Biosynthesis of Ser and Gly											
		3PG	→	Ser			*bios1*	5	0.2	5	0.2

Table 2: (continued.)

Ser	↔	Gly	+	C1	*bios2$_f$*	5	0.2	5	0.2
reversibility %						0.58	0.00	0.46	0.01
Thr	→	Gly			*bios3*	5	0.3	5	0.2

Reaction			Reaction	Zygotic Embryos		SHaM Embryos	
				Standardized w.r.t sucrose (100 moles/day)		Standardized w.r.t sucrose (100 moles/day)	
			Name	Net Flux	S.D.	Net Flux	S.D.
Fluxes towards biosynthesis (other than those related to Ser and Gly) and effluxes into medium							
CO_2	→		*resp*	**506**	**23**	**327**	**22**
G6P^c	→	biomass	*bios4*	**8**	**1**	**24**	**0.4**
G6P^p	→	biomass	*bios5*	**7**	**1**	**11**	**1**
F6P^c	→	biomass	*bios6*	1	0.2	1	0.4
P5P^c	→	biomass	*bios7*	**10**	**2**	**18**	**1**
P5P^p	→	biomass	*bios8*	1	0.1	1	0.1
E4P^p	→	biomass	*bios9*	3	0.2	2	0.4
PEP^p	→	biomass	*bios10*	6	0.3	6	0.3
ACA^p	→	biomass	*bios11*	**110**	**3**	**61**	**1**
ACA^p	→	biomass	*bios12*	3	0.2	3	0.1
Glycerol	→	biomass	*glyc*	3	1	3	2
Pyr	→	biomass	*bios13*	18	1	18	3
OAA	→	biomass	*bios14*	14	1	14	1
Glu	→	biomass	*bios15*	**11**	**1**	**8**	**1**

Supplement I: Figures and Tables

Figure S1: Comparison of proteinogenic amino acid profiles of zygotic and SHaM embryos *cv. Jack* cultured in the presence of labeled sucrose for 6 days. The media in each flask was replaced after 3 days of culture.

Figure S2: Comparison of corresponding multiplet intensities of Ala α, Phe α, Tyr α, His α, and Ser α of zygotic and SHaM embryos *cv. Jack*.

Figure S3: Simulated versus experimental comparison of: A) zygotic embryos and B) SHaM embryos.

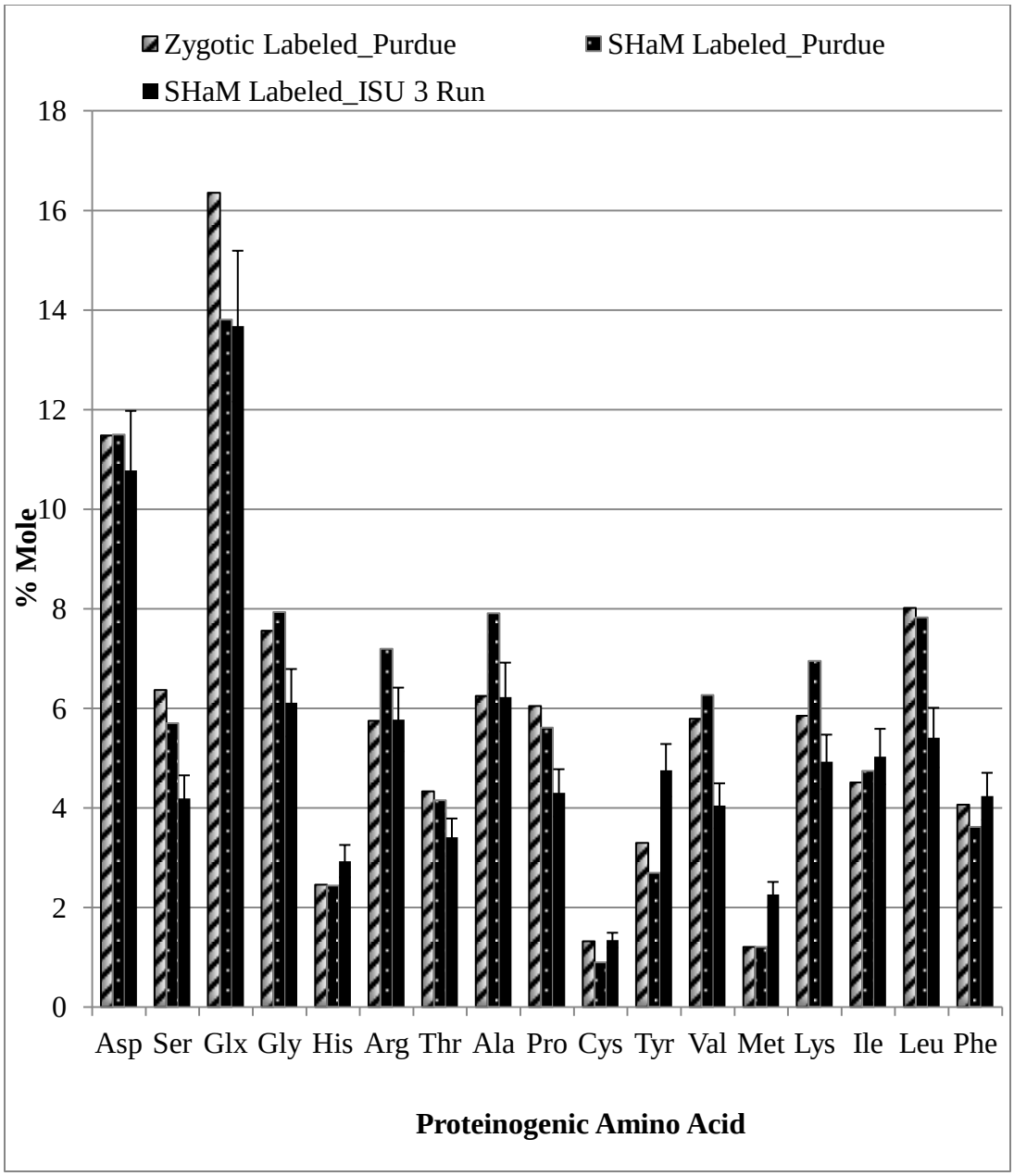
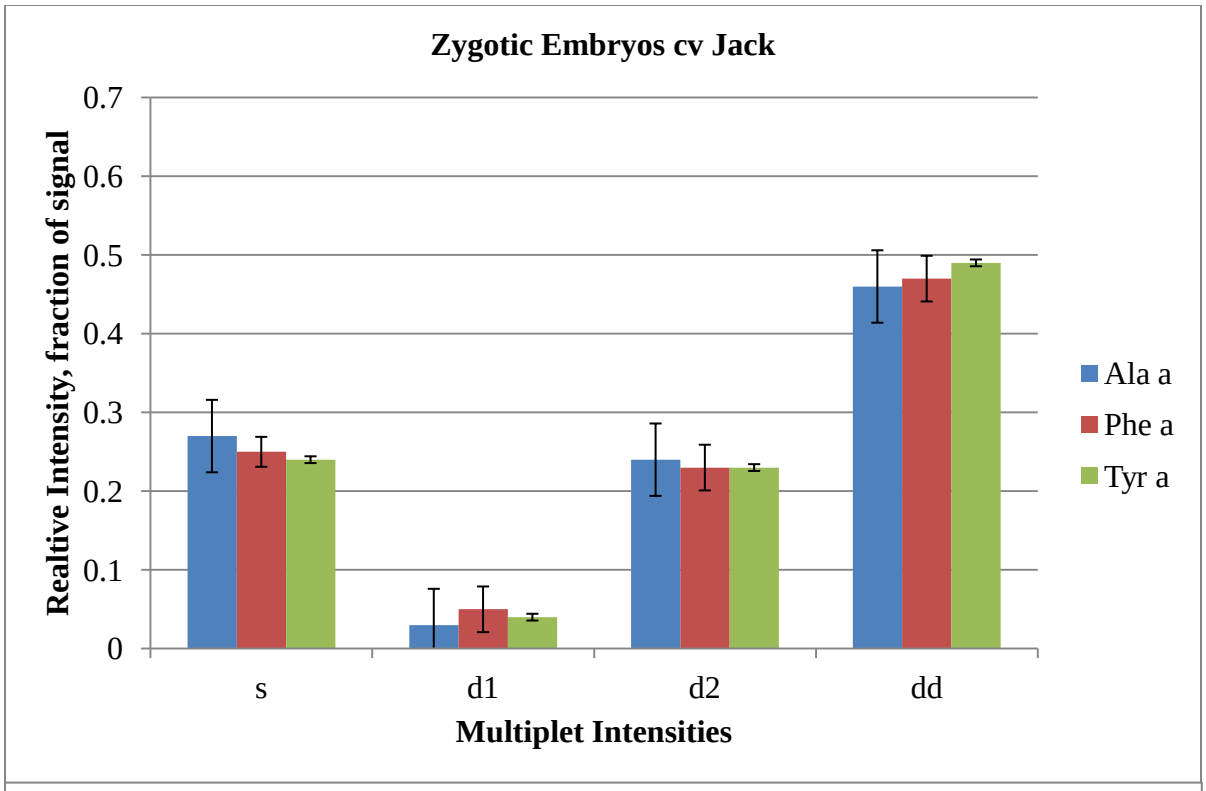

Zygotic Labeled_Purdue
SHaM Labeled_Purdue
SHaM Labeled_ISU 3 Run
18
16
14
12
10
8
6
4
2
0
% Mole
Asp Ser Glx Gly His Arg Thr Ala Pro Cys Tyr Val Met Lys Ile Leu Phe
Proteinogenic Amino Acid

Figure S1


Zygotic Embryos cv Jack
Realtive Intensity, fraction of signal
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
s
d1
d2
dd
Multiplet Intensities
Ala a
Phe a
Tyr a

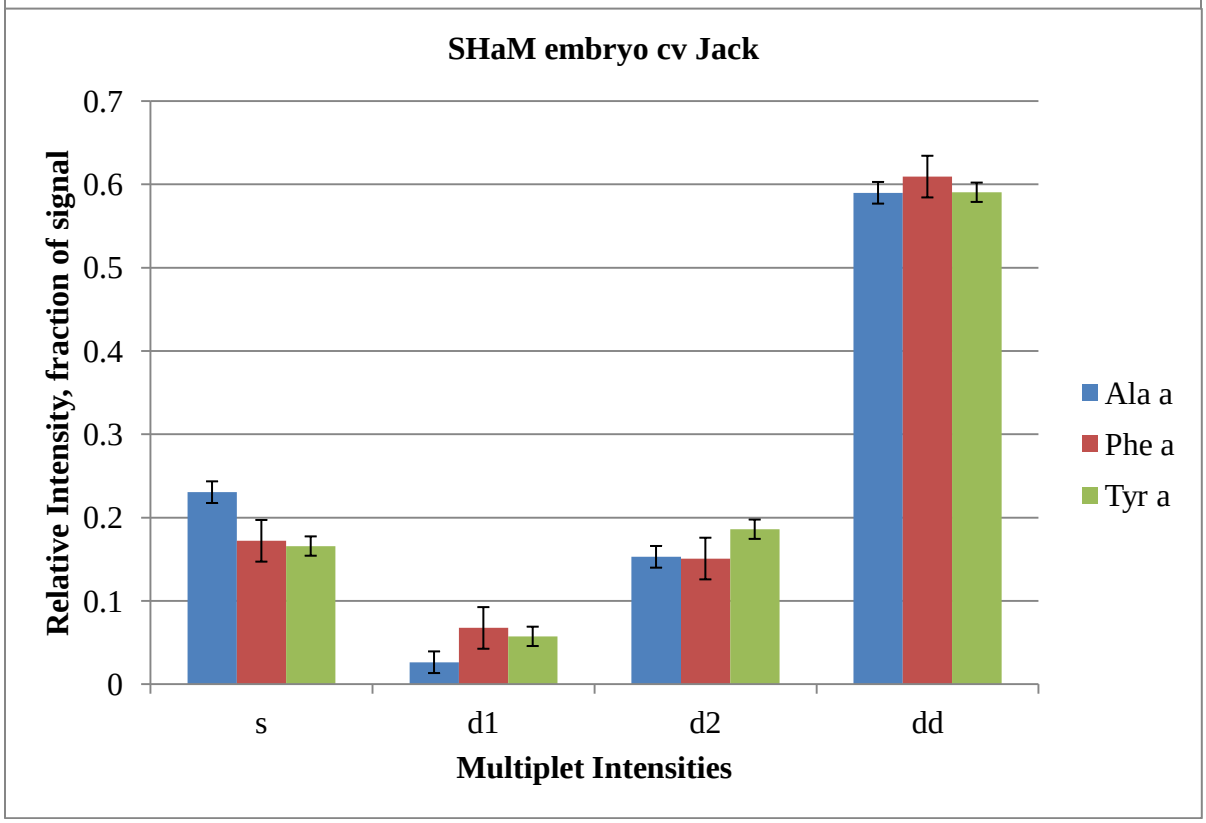

SHaM embryo cv Jack
Relative Intensity, fraction of signal
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
s
d1
d2
dd
Multiplet Intensities
Ala a
Phe a
Tyr a

Figure S2

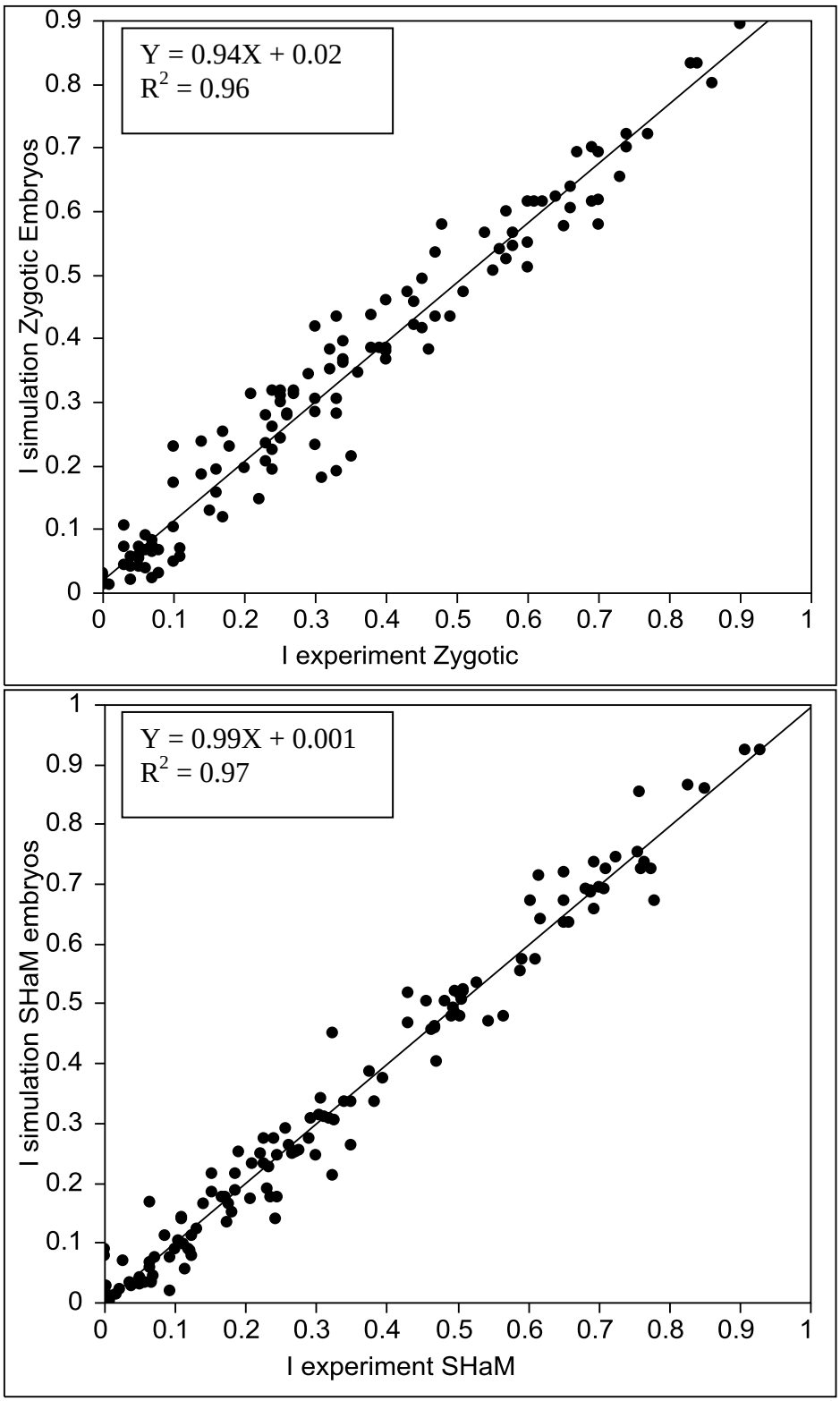

Y = 0.94X + 0.02
R2 = 0.96
I simulation Zygotic Embryos
I experiment Zygotic
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
Y = 0.99X + 0.001
R2 = 0.97
I simulation SHaM embryos
I experiment SHaM

Figure S3

Table S1: Relative multiplet intensities of amino acids from protein and glycosyl units from starch hydrolysates with their standard deviations (SDs) from 2-D [13C, 1H] HSQC spectrum (s indicates singlet, d1 and d2 indicate the first and second doublet and dd indicates the double doublet). Bold-faced carbon atom in "Isotopomer" indicates labeled carbon atom, normal font indicates unlabeled carbon atom and x indicates unknown labeling state of the carbon atom. Subscripts p, c, and m for precursor indicate the compartments plastid, cytosol, and mitochondrion in which they are present.

	Cross peak (multiplet)	Zygotic Embryos		SHaM Embryos			
		Intensity	SD	Intensity	SD	Precursor	Isotopomer
Protein	Gly α (s)	0.55	0.01	0.49	0.01	T3P/GOx	[1**2**x]
Protein	Gly α (d)	0.45	0.01	0.51	0.01	T3P/GOx	[**12**x]
Protein	Ser α (s)	0.32	0.01	0.30	0.01	Ser	[1**2**3]
Protein	Ser α (d1)	0.13	0.01	0.14	0.01	Ser	[**12**3]
Protein	Ser α (d2)	0.16	0.01	0.13	0.01	Ser	[1**23**]
Protein	Ser α (dd)	0.40	0.01	0.43	0.01	Ser	[**123**]
Protein	Ser β (s)	0.66	0.01	0.50	0.01	Ser	[x**23**]
Protein	Ser β (d)	0.34	0.01	0.50	0.01	Ser	[x**23**]
Protein	Ala α (s)	0.28	0.01	0.23	0.01	Pyr	[1**23**]
Protein	Ala α (d1)	0.03	0.01	0.03	0.01	Pyr	[**12**3]
Protein	Ala α (d2)	0.26	0.01	0.15	0.01	Pyr	[1**23**]
Protein	Ala α (dd)	0.44	0.01	0.59	0.01	Pyr	[**123**]
Protein	Ala β (s)	0.25	0.01	0.24	0.01	Pyr	[x**23**]
Protein	Ala β (d)	0.75	0.01	0.76	0.01	Pyr	[x**23**]
Protein	Ile γ2 (s)	0.29	0.01	0.24	0.01	Pyr_p	[x**23**]
Protein	Ile γ2 (d)	0.71	0.01	0.76	0.01	Pyr_p	[x**23**]
Protein	Leu δ1 (s)	0.34	0.01	0.29	0.01	Pyr_p	[x**23**]
Protein	Leu δ1 (d)	0.66	0.01	0.71	0.01	Pyr_p	[x**23**]
Protein	Val α (s)	0.41	0.01	0.38	0.01	Pyr_p	[1**2**x]·[x2x]

Table S1: (continued.)

Protein	Val α (d1)	0.44	0.01	0.53	0.01	Pyr_p	[**12**x]·[x2x]
Protein	Val α (d2)	0.04	0.01	0.05	0.01	Pyr_p	[1**2**x]·[x**2**x]
Protein	Val α (dd)	0.10	0.01	0.07	0.01	Pyr_p	[**12**x]·[x**2**x]
Protein	Val γ1 (s)	0.28	0.01	0.23	0.01	Pyr_p	[x2**3**]
Protein	Val γ1 (d)	0.72	0.01	0.78	0.01	Pyr_p	[x**23**]
Protein	Val γ2 (s)	0.90	0.01	0.91	0.01	Pyr_p	[x2x]·[xx**3**]
Protein	Val γ2 (d)	0.10	0.01	0.09	0.01	Pyr_p	[x**2**x]·[xx**3**]

	Cross peak (multiplet)	**Zygotic Embryos**		**SHaM Embryos**			
		Intensity	**SD**	**Intensity**	**SD**	**Precursor**	**Isotopomer**
Protein	Leu δ2 (s)	0.91	0.01	0.93	0.01	Pyr_p	[x2x]·[xx**3**]
Protein	Leu δ2 (d)	0.09	0.01	0.07	0.01	Pyr_p	[x**2**x]·[xx**3**]
Protein	Phe α (s)	0.25	0.01	0.17	0.03	PEP_p	[1**2**3]
Protein	Phe α (d1)	0.05	0.01	0.07	0.03	PEP_p	[**12**3]
Protein	Phe α (d2)	0.23	0.01	0.15	0.03	PEP_p	[1**23**]
Protein	Phe α (dd)	0.47	0.01	0.61	0.03	PEP_p	[**123**]
Protein	Phe β (s)	0.21	0.01	0.25	0.01	PEP_p	[x**23**]·[x2x]
Protein	Phe β (d1)	0.65	0.01	0.62	0.01	PEP_p	[x**23**]·[x2x]
Protein	Phe β (d2)	0.06	0.01	0.02	0.01	PEP_p	[x**23**]·[x**2**x]
Protein	Phe β (dd)	0.07	0.01	0.12	0.01	PEP_p	[x**23**]·[x**2**x]
Protein	Tyr α (s)	0.24	0.01	0.17	0.01	PEP_p	[1**2**3]
Protein	Tyr α (d1)	0.05	0.01	0.06	0.01	PEP_p	[**12**3]
Protein	Tyr α (d2)	0.23	0.01	0.19	0.01	PEP_p	[1**23**]
Protein	Tyr α (dd)	0.49	0.01	0.59	0.01	PEP_p	[**123**]

Table S1: (continued.)

Protein	Tyr β (s)	0.19	0.01	0.24	0.01	PEP_p	[x2**3**]·[x2x]
Protein	Tyr β (d)	0.69	0.01	0.76	0.01	PEP_p	[x**23**]·[x2x] + [x23]·[x2x]
Protein	Tyr β (t)	0.12	0.01	0.00	0.01	PEP_p	[x**23**]·[x**2**x]
Protein	Leu α (s)	0.29	0.01	0.19	0.02	$ACoA_p$/Pyr_p	[1**2**]·[x2x]
Protein	Leu α (d1)	0.60	0.01	0.60	0.02	$ACoA_p$/Pyr_p	[**12**]·[x2x]
Protein	Leu α (d2)	0.04	0.01	0.09	0.02	$ACoA_p$/Pyr_p	[1**2**]·[x**2**x]
Protein	Leu α (dd)	0.07	0.01	0.11	0.02	$ACoA_p$/Pyr_p	[**12**]·[x**2**x]
Protein	Leu b (s)	0.86	0.01	0.76	0.02	$ACoA_p$/Pyr_p	[x2]·[x**2**x].[x2x]
Protein	Leu b (d)	0.14	0.01	0.24	0.02	$ACoA_p$/Pyr_p	[x**2**]·[x**2**x].[x2x] +[x2]·[x2x].[x2x]
Protein	Leu b (t)	0.00	0.01	0.00	0.02	$ACoA_p$/Pyr_p	[x**2**]·[x**2**x].[x**2**x]
Protein	His β (s)	0.33	0.01	0.30	0.01	$P5P_p$	[x23**4**x]
Protein	His β (d1)	0.04	0.01	0.068	0.01	$P5P_p$	[x**23**4x]
Protein	His β (d2)	0.50	0.01	0.57	0.01	$P5P_p$	[x2**34**x]
Protein	His β (dd)	0.13	0.01	0.06	0.01	$P5P_p$	[x**234**x]

	Cross peak (multiplet)	**Zygotic Embryos**		**SHaM Embryos**			
		Intensity	**SD**	**Intensity**	**SD**	**Precursor**	**Isotopomer**
Protein	His δ2 (s)	0.66	0.01	0.69	0.08	$P5P_p$	[**1**2xxx]
Protein	His δ2 (d)	0.34	0.01	0.31	0.08	$P5P_p$	[**12**xxx]
Protein	Tyr δ1 (s)	0.27	0.01	0.21	0.01	PEP_p/$E4P_p$	[x**23**]·[1xxx]
Protein	Tyr δ1 (d)	0.69	0.01	0.69	0.01	PEP_p/$E4P_p$	[x**23**]·[1xxx] + [x**23**]·[**1**xxx]
Protein	Tyr δ1 (t)	0.03	0.01	0.10	0.01	PEP_p/$E4P_p$	[x**23**]·[**1**xxx]
Protein	Tyr ε1 (s)	0.58	0.01	0.47	0.01	PEP_p/$E4P_p$	[xx3]·[**1**2xx]

Table S1: (continued)

Protein	Tyr ε1 (d)	0.24	0.01	0.21	0.01	PEP_p/$E4P_p$	[xx**3**]·[**1**2xx] + [xx3]·[**12**xx]
Protein	Tyr ε1 (t)	0.18	0.01	0.32	0.01	PEP_p/$E4P_p$	[xx**3**]·[**12**xx]
Protein	Arg β (s)	0.84	0.01	0.65	0.02	Glu_p	[x2**3**4x]
Protein	Arg β (d)	0.16	0.01	0.35	0.02	Glu_p	[x**23**4x] + [x2**34**x]
Protein	Arg β (t)	0.00	0.01	0.00	0.02	Glu_p	[x**234**x]
Protein	Arg δ (s)	0.74	0.01	0.51	0.01	Glu_p	[xxx4**5**]
Protein	Arg δ (d)	0.26	0.01	0.49	0.01	Glu_p	[xxx**45**]
Protein	Glu β (s)	0.79	0.01	0.66	0.04	Glu	[x2**3**4x]
Protein	Glu β (d)	0.21	0.01	0.34	0.04	Glu	[x**23**4x] + [x2**34**x]
Protein	Glu β (t)	0.00	0.01	0.01	0.04	Glu	[x**234**x]
Protein	Glu γ (s)	0.69	0.01	0.48	0.03	Glu	[xx3**4**5]
Protein	Glu γ (d1)	0.04	0.01	0.02	0.03	Glu	[xx**34**5]
Protein	Glu γ (d2)	0.24	0.01	0.46	0.03	Glu	[xx3**45**]
Protein	Glu γ (dd)	0.04	0.01	0.04	0.03	Glu	[xx**345**]
Protein	Pro γ (s)	0.60	0.01	0.46	0.02	Glu_p	[xx3**4**5]
Protein	Pro γ (d)	0.40	0.01	0.54	0.02	Glu_p	[xx**34**5] + [xx3**45**]
Protein	Pro γ (t)	0.00	0.01	0.00	0.02	Glu_p	[xx**345**]
Protein	Pro δ (s)	0.77	0.01	0.50	0.01	Glu_p	[xxx4**5**]
Protein	Pro δ (d)	0.23	0.01	0.50	0.01	Glu_p	[xxx**45**]
Protein	Asp α (s)	0.33	0.01	0.27	0.02	OAA	[1**2**3x]
Protein	Asp α (d1)	0.04	0.01	0.11	0.02	OAA	[**12**3x]
Protein	Asp α (d2)	0.41	0.01	0.23	0.02	OAA	[1**23**x]
Protein	Asp α (dd)	0.22	0.01	0.40	0.02	OAA	[**123**x]
Protein	Asp β (s)	0.38	0.01	0.26	0.01	OAA	[x2**3**4]

Table S1: (continued.)

	Cross peak	Zygotic Embryos		SHaM Embryos			
	(multiplet)	Intensity	SD	Intensity	SD	Precursor	Isotopomer
Protein	Asp β (d1)	0.37	0.01	0.47	0.01	OAA	[x**2**3**4**]
Protein	Asp β (d2)	0.10	0.01	0.09	0.01	OAA	[x2**34**]
Protein	Asp β (dd)	0.15	0.01	0.18	0.01	OAA	[x**234**]
Protein	Ile α (s)	0.71	0.01	0.51	0.02	OAA_p/Pyr_p	[1**2**xx]·[x2x]
Protein	Ile α (d1)	0.22	0.01	0.24	0.02	OAA_p/Pyr_p	[**12**xx]·[x2x]
Protein	Ile α (d2)	0.05	0.01	0.08	0.02	OAA_p/Pyr_p	[1**2**xx]·[x**2**x]
Protein	Ile α (dd)	0.03	0.01	0.16	0.02	OAA_p/Pyr_p	[**12**xx]·[x**2**x]
Protein	Ile γ1(s)	0.73	0.01	0.65	0.02	Pyr_p/OAA_p	[x2x]·[xx**3**4]
Protein	Ile γ1(d)	0.27	0.01	0.35	0.02	Pyr_p/OAA_p	[x**2**x]·[xx**3**4] + [x2x]·[xx**34**]
Protein	Ile γ1(t)	0.00	0.01	0.00	0.02	Pyr_p/OAA_p	[x**2**x]·[xx**34**]
Protein	Ile δ (s)	0.80	0.01	0.68	0.02	OAA_p	[xx3**4**]
Protein	Ile δ (d)	0.20	0.01	0.32	0.02	OAA_p	[xx**34**]
Protein	Thr α (s)	0.40	0.01	0.23	0.06	OAA_p	[1**2**3x]
Protein	Thr α (d1)	0.10	0.01	0.12	0.06	OAA_p	[**12**3x]
Protein	Thr α (d2)	0.33	0.01	0.32	0.06	OAA_p	[1**23**x]
Protein	Thr α (dd)	0.18	0.01	0.32	0.06	OAA_p	[**123**x]
Protein	Thr γ2 (s)	0.41	0.01	0.71	0.01	OAA_p	[xx3**4**]
Protein	Thr γ2 (d)	0.59	0.01	0.29	0.01	OAA_p	[xx**34**]
Protein	Lys γ (s)	0.74	0.01	0.62	0.02	OAA_p/Pyr_p	[xx3**4**]·[xx3]
Protein	Lys γ (d)	0.26	0.01	0.38	0.02	OAA_p/Pyr_p	[xx**34**]·[xx3] + [xx3**4**]·[xx**3**]
Protein	Lys γ (t)	0.00	0.01	0.00	0.02	OAA_p/Pyr_p	[xx**34**]·[xx**3**]
Protein	Lys β (s)	0.24	0.01	0.22	0.03	OAA_p/Pyr_p	½{[x2**3**4] +

Table S1: (continued.)

							[x**23**]·[xxx4]}c
Protein	Lys β (d)	0.68	0.01	0.78	0.03	OAA_p/Pyr_p	½{[x**23**4] + [x2**34**] + [x**23**]·[xxx4]+[x23]·[xxx**4**]
Protein	Lys β (t)	0.07	0.01	0.00	0.03	OAA_p/Pyr_p	½{[x**234**] + [x**23**]·[xxx**4**]}
Protein	Lys δ (s)	0.45	0.01	0.22	0.01	OAA_p/Pyr_p	½{[x23**4**] + [x2**3**]·[xxx4]}
Protein	Lys δ (d)	0.49	0.01	0.61	0.01	OAA_p/Pyr_p	½{[x**23**4] + [x2**34**] + [x**23**]·[xxx4] + [x2**3**]·[xxx**4**]}

	Cross peak	Zygotic Embryos		SHaM Embryos			
	(multiplet)	Intensity	SD	Intensity	SD	Precursor	Isotopomer
Protein	Lys δ (t)	0.061	0.01	0.17	0.01	OAA_p/Pyr_p	½{[x**234**] + [x**23**]·[xxx**4**]}
Protein	Lys ε (s)	0.34	0.01	0.31	0.01	OAA_p/Pyr_p	½{[x**2**3] + [x**2**3x]}
Protein	Lys ε (d)	0.66	0.01	0.69	0.01	OAA_p/Pyr_p	½{[x**23**] + [x**23**x]}
Protein	HMF 1 (s)	0.63	0.02	0.63	0.01	$G6P_c/F6P_c$	[**1**2xxxx]
Protein	HMF 1 (d)	0.37	0.02	0.37	0.01	$G6P_c/F6P_c$	[**12**xxxx]
Protein	LVA 3 (s)	0.18	0.01	0.13	0.01	$G6P_c/F6P_c$	[x2**3**4xx]
Protein	LVA 3 (d1)	0.03	0.01	0.12	0.01	$G6P_c/F6P_c$	[x**23**4xx]
Protein	LVA 3 (d2)	0.01	0.01	0.01	0.01	$G6P_c/F6P_c$	[x2**34**xx]
Protein	LVA 3 (dd)	0.77	0.01	0.74	0.01	$G6P_c/F6P_c$	[x**234**xx]
Protein	LVA 6 (s)	0.10	0.01	0.10	0.01	$G6P_c/F6P_c$	[xxx45**6**]
Protein	LVA 6 (d1)	0.07	0.01	0.09	0.01	$G6P_c/F6P_c$	[xxx4**56**]
Protein	LVA 6 (d2)	0.01	0.01	0.06	0.01	$G6P_c/F6P_c$	[xxx**45**6]
Protein	LVA 6 (dd)	0.83	0.01	0.75	0.01	$G6P_c/F6P_c$	[xxx**456**]

Table S1: (continued.)

Starch	HMF 1 (s)	0.43	0.03	0.28	0.07	$G6P_p$	[**1**2xxxx]
Starch	HMF 1 (d)	0.57	0.03	0.72	0.07	$G6P_p$	[**12**xxxx]
Starch	LVA 3 (s)	0.47	0.01	0.41	0.02	$G6P_p$	[x2**3**4xx]
Starch	LVA 3 (d1)	0.20	0.01	0.21	0.02	$G6P_p$	[x**23**4xx]
Starch	LVA 3 (d2)	0.07	0.01	0.02	0.02	$G6P_p$	[x2**34**xx]
Starch	LVA 3 (dd)	0.26	0.01	0.36	0.02	$G6P_p$	[x**234**xx]
Starch	LVA 6 (s)	0.25	0.01	0.29	0.01	$G6P_p$	[xxx45**6**]
Starch	LVA 6 (d1)	0.08	0.01	0.09	0.01	$G6P_p$	[xxx4**56**]
Starch	LVA 6 (d2)	0.04	0.01	0.01	0.01	$G6P_p$	[xxx**4**5**6**]
Starch	LVA 6 (dd)	0.61	0.01	0.62	0.01	$G6P_p$	[xxx**456**]

Table S2: Values of absolute fluxes for the two embryo types analyzed in this study with the unit of umol metabolite/gDwt-Day. Values show best fit of flux estimation with standard deviation as determined based on Monte Carlo simulation of 500 times.

Reaction	Reaction Name	Zygotic Embryos umol metabolite/g Dwt-day		SHaM Embryos umol metabolite/g Dwt-day	
		Net Flux	S.D.	Net Flux	S.D.
Glycolysis and oxPPP					
$G6P^c \leftrightarrow F6P^c$	hxi_f^c	-1223	349	-841	118
reversibility %		0.58	0.22	0.59	0.27
$G6P^p \leftrightarrow F6P^p$	hxi_f^p	-900	146	-734	134
reversibility %		0.87	0.11	0.73	0.15
$G6P^c \rightarrow P5P^c + CO2$	pgl^c	1562	420	1137	206
$G6P^p \rightarrow P5P^p + CO2$	pgl^p	1879	244	1244	243
$P5P^c + P5P^c \leftrightarrow S7P^c + T3P^c$	$tktA_f^c$	128	63	57	24
reversibility %		0.87	0.14	0.81	0.08
$S7P^c + T3P^c \leftrightarrow F6P^c + E4P^c$	tal_f^c	128	63	57	24
reversibility %		0.99	0.01	0.88	0.10
$P5P^c + E4P^c \leftrightarrow F6P^c + T3P^c$	$tktB_f^c$	128	63	57	24
reversibility %		0.58	0.20	0.37	0.14
$P5P^p + P5P^p \leftrightarrow S7P^p + T3P^p$	$tktA_f^p$	736	143	675	60
reversibility %		0.97	0.04	0.56	0.43
$S7P^p + T3P^p \leftrightarrow F6P^p + E4P^p$	tal_f^p	736	143	675	60
reversibility %		0.83	0.05	0.55	0.23
$P5P^p + E4P^p \leftrightarrow F6P^p + T3P^p$	$tktB_f^p$	698	143	647	61
reversibility %		0.31	0.21	0.77	0.16
$F6P^c \rightarrow T3P^c + T3P^c$	pfk^c	174	72	250	95
$F6P^p \rightarrow T3P^p + T3P^p$	pfk^p	1184	68	374	128
$T3P^p + T3P^p \rightarrow F6P^p$	$f16bp^p$	2884	358	1448	308
$T3P \rightarrow 3PG$	gap	3677	130	1772	190
$3PG \rightarrow PEP$	eno	3608	128	1742	188
$PEP \rightarrow Pyr$	pyk	2470	79	1012	110
$Pyr^p \rightarrow ACA^p + CO_2$	pdh^p	1527	58	579	67
$Pyr^m \rightarrow ACA^m + CO_2$	pdh^m	799	46	611	40
$Pyr^c \rightarrow Pyr^m$	$pyrT$	934	105	682	33
$G6P^c \leftrightarrow G6P^p$	$g6pT_f$	665	301	202	110
reversibility %		0.43	0.08	0.50	0.25
$P5P^c \leftrightarrow P5P^p$	$p5pT_f$	303	268	628	185
reversibility %		0.87	0.17	0.81	0.19
$P5P^p + CO_2 \rightarrow 3PG + 3PG$	$rubisco$	125	65	64	28

Table S2: (continued.)

		$T3P^c$	↔	$T3P^p$			$t3pT_f$	672	153	185	217
		reversibility %						0.93	0.03	0.31	0.28
Reaction							Reaction	Zygotic Embryos		SHaM Embryos	
							Name	umol metabolite/g Dwt-day		umol metabolite/g Dwt-day	
								Net Flux	S.D.	Net Flux	S.D.
TCA cycle											
ACA^m	+	OAA^m	→	$ICit^m$			$csaco^m$	763	116	312	49
		$ICit^m$	→	aKG^m	+	CO_2	$icdh^m$	763	116	312	49
		aKG^m	→	Scn^m	+	CO_2	$akgdh^m$	634	122	266	49
		Scn^m	→	Mal^m			$sdh_{f1}{}^m$	634	122	266	49
		Scn^m	↔	Mal^m			$sdh_{f2}{}^m$	634	122	883	112
		reversibility %						0.85	0.04	0.39	0.03
		Mal^m	↔	OAA^m			$mdh_f{}^m$	795	115	341	49
		reversibility %						0.49	0.35	0.54	0.31
Anaplerotic reactions											
PEP^c	+	CO_2	→	OAA^c			ppc^c	377	81	154	14
		Mal^m	→	Pyr^m	+	CO_2	me^m	136	45	62	13
		Mal^p	→	Pyr^p	+	CO_2	me^p	81	16	41	12
Substrate entry											
		Suc^{ext}	→	$G6P^c$	+	$F6P^c$	*subs1*	1373	152	960	33
		Gln^{ext}	↔	Gln^p			$subs2_f$	763	31	478	4
		reversibility %						0.99	0.00	0.90	0.01
Glutamate assimilation											
		Gln^p	→	Glu^p			as_f	802	26	312	49
		Glu^p	→	aKG^p			$gdh_f{}^p$	725	39	312	49
aKG^p	+	Gln^p	→	Glu^p	+	Glu^p	$gogat^p$	236	66	266	49
Malate shuttle											
		Mal^m	↔	Mal^c			$malT1_f$	-116	14	-24	21
		reversibility %						0.99	0.00	0.95	0.02
		Mal^c	→	Mal^p			*malT2*	493	67	179	36
		Mal^c	→	OAA^c			mdh^c	377	81	154	14
		Mal^p	→	OAA^P			mdh^p	180	9	131	6
Biosynthesis of Ser and Gly											
		3PG	→	Ser			*bios1*	69	3	50	2
		Ser	↔	Gly	+	C1	$bios2_f$	69	3	50	2
		reversibility %						0.58	0.00	0.46	0.01

Table S2: (continued.)

Thr	→	Gly	*bios3*	72	4	51	2
Reaction			Reaction	Zygotic Embryos		SHaM Embryos	
			Name	umol metabolite/g Dwt-day		umol metabolite/g Dwt-day	
				Net Flux	S.D.	Net Flux	S.D.
Fluxes towards biosynthesis (other than those related to Ser and Gly) and effluxes into medium							
CO_2	→		*resp*	6951	316	3139	207
$G6P^c$	→	biomass	*bios4*	111	11	233	3
$G6P^p$	→	biomass	*bios5*	94	9	108	7
$F6P^c$	→	biomass	*bios6*	8	2	12	3
$P5P^c$	→	biomass	*bios7*	134	28	176	5
$P5P^p$	→	biomass	*bios8*	12	0	7	1
$E4P^p$	→	biomass	*bios9*	39	3	18	4
PEP^p	→	biomass	*bios10*	78	4	55	3
ACA^p	→	biomass	*bios11*	1514	36	587	11
ACA^p	→	biomass	*bios12*	41	2	27	1
Glycerol	→	biomass	*glyc*	45	15	30	14
Pyr	→	biomass	*bios13*	245	9	170	26
OAA	→	biomass	*bios14*	180	9	131	6
Glu	→	biomass	*bios15*	144	13	74	5

Table S3: Relative fluxes are standardized with respect to sucrose 100 moles/day, and are expressed as such. Comparison between zygotic embryos *cv. Jack* with *cv.* Evans (*these values has been published previously (Iyer *et al.*, 2008). Each flux is represented by the name of the gene encoding the enzyme catalyzing the metabolic reaction subscripts f indicates forward reaction and superscripts c, p an dm show cytosol, plastid and mitochondrion, respectively.

Reaction	Reaction	Zygotic Embryos cv. Jack		Zygotic Embryos cv. Evans*	
	Name	Standardized w.r.t sucrose (100 moles/day)		Standardized w.r.t sucrose (100 moles/day)	
		Net Flux	S.D.	Net Flux	S.D.
Glycolysis and oxPPP					
$G6P^c \leftrightarrow F6P^c$	hxi_f^c	-89	25	-100	16
reversibility %		0.58	0.22	0.52	0.27
$G6P^p \leftrightarrow F6P^p$	hxi_f^p	-66	11	-38	44
reversibility %		0.87	0.11	0.97	0.03
$G6P^c \rightarrow P5P^c + CO2$	pgl^c	111	31	95	40
$G6P^p \rightarrow P5P^p + CO2$	pgl^p	137	18	114	52
$P5P^c + P5P^c \leftrightarrow S7P^c + T3P^c$	$tktA_f^c$	9	5	6	3
reversibility %		0.87	0.14	0.90	0.18
$S7P^c + T3P^c \leftrightarrow F6P^c + E4P^c$	tal_f^c	9	5	6	3
reversibility %		0.99	0.01	0.47	0.25
$P5P^c + E4P^c \leftrightarrow F6P^c + T3P^c$	$tktB_f^c$	9	5	6	3
reversibility %		0.58	0.20	0.67	0.11
$P5P^p + P5P^p \leftrightarrow S7P^p + T3P^p$	$tktA_f^p$	54	10	57	16
reversibility %		0.97	0.04	0.89	0.14
$S7P^p + T3P^p \leftrightarrow F6P^p + E4P^p$	tal_f^p	54	10	57	16
reversibility %		0.83	0.05	0.76	0.16
$P5P^p + E4P^p \leftrightarrow F6P^p + T3P^p$	$tktB_f^p$	51	10	53	16
reversibility %		0.31	0.21	0.34	0.31
$F6P^c \rightarrow T3P^c + T3P^c$	pfk^c	13	5	13	15
$F6P^p \rightarrow T3P^p + T3P^p$	pfk^p	86	5	72	18
$T3P^p + T3P^p \rightarrow F6P^p$	$f16bp^p$	210	26	214	61
$T3P \rightarrow 3PG$	gap	268	9	223	18
$3PG \rightarrow PEP$	eno	263	9	218	17
$PEP \rightarrow Pyr$	pyk	180	6	193	15
$Pyr^p \rightarrow ACA^p + CO_2$	pdh^p	111	4	120	1
$Pyr^m \rightarrow ACA^m + CO_2$	pdh^m	58	3	58	14
$Pyr^c \rightarrow Pyr^m$	$pyrT$	68	7	58	14
$G6P^c \leftrightarrow G6P^p$	$g6pT_f$	48	22	84	40
reversibility %		0.43	0.08	0.50	0.27
$P5P^c \leftrightarrow P5P^p$	$p5pT_f$	22	20	54	41
reversibility %		0.09	0.17	0.76	0.29
$P5P^p + CO_2 \rightarrow 3PG + 3PG$	$rubisco$	9	5	N/A	N/A
$T3P^c \leftrightarrow T3P^p$	$t3pT_f$	49	11	46	30
reversibility %		0.93	0.03	0.96	0.07

Table S3: (continued.)

Reaction								Reaction	Zygotic Embryos cv. Jack		Zygotic Embryos cv. Evans*	
								Name	Standardized w.r.t sucrose (100 moles/day)		Standardized w.r.t sucrose (100 moles/day)	
									Net Flux	S.D.	Net Flux	S.D.
TCA cycle												
ACA^{m}	+	OAA^{m}	→	ICit^{m}				*csaco^{m}*	56	8	53	14
		ICit^{m}	→	aKG^{m}	+	CO_2		*icdh^{m}*	56	8	48	14
		aKG^{m}	→	Scn^{m}	+	CO_2		*akgdh^{m}*	46	9	6	21
		Scn^{m}	→	Mal^{m}				*sdh$_{f1}$^{m}*	46	9	35	10
		Scn^{m}	↔	Mal^{m}				*sdh$_{f2}$^{m}*	46	9	67	14
		reversibility %							0.85	0.04	0.89	0.04
		Mal^{m}	↔	OAA^{m}				*mdh$_{f}$^{m}*	58	8	66	14
		reversibility %							0.49	0.35	0.45	0.32
Anaplerotic reactions												
PEP^{c}	+	CO_2	→	OAA^{c}				*ppc^{c}*	27	6	22	3
		Mal^{m}	→	Pyr^{m}	+	CO_2		*me^{m}*	10	3	9	4
		Mal^{p}	→	Pyr^{p}	+	CO_2		*me^{p}*	6	1	6	3
Substrate entry												
		Suc^{ext}	→	G6P^{c}	+	F6P^{c}		*subs1*	100	0	100	0
		Gln^{ext}	↔	Gln^{p}				*subs2$_{f}$*	65	1	50	2
		reversibility %							0.99	0.00	0.84	0.02
Glutamate assimilation												
		Gln^{p}	→	Glu^{p}				*as$_{f}$*	58	2	872	457
		Glu^{p}	→	aKG^{p}				*gdh$_{f}$^{p}*	53	3	875	457
aKG^{p}	+	Gln^{p}	→	Glu^{p}	+	Glu^{p}		*gogat^{p}*	17	5	924	460
Malate shuttle												
		Mal^{m}	↔	Mal^{c}				*malT1$_{f}$*	-9	1	-8	5
		reversibility %							0.99	0.01	0.97	0.01
		Mal^{c}	→	Mal^{p}				*malT2*	36	5	14	8
		Mal^{c}	→	OAA^{c}				*mdh^{c}*	27	6	22	3
		Mal^{p}	→	OAA^{P}				*mdh^{p}*	13	1	8	5
Biosynthesis of Ser and Gly												
		3PG	→	Ser				*bios1*	5	0.2	5	0.3
		Ser	↔	Gly	+	C1		*bios2$_{f}$*	5	0.2	5	0.3
		reversibility %							.58	0.00	0.41	0.01
		Thr	→	Gly				*bios3*	5	0.3	5	0.2

Table S3: (continued.)

			Name	Standardized w.r.t sucrose (100 moles/day)		Standardized w.r.t sucrose (100 moles/day)	
				Net Flux	S.D.	Net Flux	S.D.
Fluxes towards biosynthesis (other than those related to Ser and Gly) and effluxes into medium							
CO_2	→		*resp*	506	23	519	24
$G6P^c$	→	biomass	*bios4*	8	1	21	3
$G6P^p$	→	biomass	*bios5*	7	1	9	0.3
$F6P^c$	→	biomass	*bios6*	1	0.2	0.4	0.3
$P5P^c$	→	biomass	*bios7*	10	2	22	1
$P5P^p$	→	biomass	*bios8*	1	0.1	1	0.01
$E4P^p$	→	biomass	*bios9*	3	0.2	4	0.1
PEP^p	→	biomass	*bios10*	6	0.3	8	0.1
ACA^p	→	biomass	*bios11*	110	3	114	1
ACA^p	→	biomass	*bios12*	3	0.2	6	0.1
Glycerol	→	biomass	*glyc*	3	1	N/A	N/A
Pyr	→	biomass	*bios13*	18	1	29	0.3
OAA	→	biomass	*bios14*	13	1	8	0.1
Glu	→	biomass	*bios15*	11	1	17	0.2

Supplement II: Detailed Calculations

I. <u>Calculation of the sucrose and glutamine consumption rate: Somatic Embryos</u>

The culture media were measured with HPLC to obtain sucrose and glutamine concentration. Two approximation techniques were employed for estimating the substrates consumption rates. First, polynomial fit which involves the approximation of total sucrose concentration profile by polynomial function. Since specific growth rate of somatic embryos is assumed to be constant, can be described by plotting equation (1), the rate of sucrose consumed, against biomass, equation (2) can be used to calculated specific sucrose consumption rate.

$$\frac{dx}{dt} = k \tag{1}$$

$$\frac{dC_{subs}}{dt} = r_{subs}x \tag{2}$$

Table 1D: Substrate consumption rate data from somatic embryos

Culture day (day)	Sucrose (mM)	Glucose (mM)	Total Sucrose (mM)	Biomass Accumulation DW (g)	dC/dt (mmol/day)
0	82	5	84	0	-0.2
1	76	5	78	0.1	**-0.4**
2	65	4	67	0.2	**-0.5**
3	46	2	47	0.3	**-0.5**
5	28	0	28	0.5	-0.4
7	16	0	16	0.6	-0.1

Table 2D: Compare two different techniques of calculate sucrose consumption rate

Polynomial fit (mmol/g Dwt day)	-0.6
End point calculation (mmol/g DW day)	-0.5

Another way to calculate the sucrose consumption rate is by calculating the difference in sucrose concentration between day 7 and day zero of the culture as shown in equation (3).

$$\frac{(C_{day7} - C_{dayo}) * V}{(t_{final} - t_{initial}) * (x_{day7} - x_{dayo})} \quad (3)$$

II. <u>Calculation of external fluxes contributing towards protein and lipid synthesis</u>

The external fluxes were determined from the amino acid HPLC analysis, coupled with the precursor-amino acid stoichiometry, shows in Table 3D. OAA is a precursor for Asx, Ile, Lys, Met and Thr. The biochemistry of the pathways indicate that 1 mole of Asx, Ile, Met and Thr each requires 1 mole OAA. 1 mole of Lys, however, requires 2 moles of OAA. Similar analysis could be applied to the other fluxes as well. Table 4D shows the relationship of the stoichiometry between metabolites and corresponding amino acids. The calculation of moles of lipids in the system was carried out as follows. The lipid percentage was measured gravimetrically as described in Methods section. Tristearin was assumed to be the major lipid present and the molecular weight of tristearin (tri-18C) fatty acids was considered for the calculations ($C_{57}H_{110}O_6$ molecular weight = 891.48 g/mole).

$$27\ \text{A-CoA} + 1\ \text{T3P} \rightarrow 1\ \text{Lipid}$$

Thus, the corresponding amount of A-CoA and T3P can be calculated.

Table 3D: External fluxes to be input to the NMR2Flux program

External Fluxes	Precursor metabolite → Amino Acids/Lipids
1	A-CoA → Lipids
2	P5P → His + Nucleotides
3	E4P → Phe + Tyr
4	3-PG → Gly + Ser
5	PEP → Phe + Tyr
6	Pyr → Ala + Val + Ile + Lys + Leu
7	A-CoA → Leu
8	OAA → Asx + Ile + Lys + Met + Thr
9	Glu → Glx + Pro + Arg

Table 4D: Molar ratios of metabolites to corresponding amino acids

Precursor Metabolite → Amino Acid ↓	3-PGA	Pyr	PEP	E4P	A-CoA	OAA	2-OxoG	R5P
Ala	0	1	0	0	0	0	0	0
Asx	0	0	0	0	0	1	0	0
Glx	0	0	0	0	0	0	1	0
Phe	0	0	2	1	0	0	0	0
Gly	1	0	0	0	0	0	0	0
His	0	0	0	0	0	0	0	1
Ile	0	1	0	0	0	1	0	0
Lys (50%)	0	1	0	0	0	1	0	0
Lys (50%)	0	1	0	0	0	1	0	0
Leu	0	2	0	0	1	0	0	0
Met	0	0	0	0	0	1	0	0
Pro	0	0	0	0	0	0	1	0
Arg	0	0	0	0	0	0	1	0
Ser	1	0	0	0	0	0	0	0
Thr	0	0	0	0	0	1	0	0
Val	0	2	0	0	0	0	0	0
Tyr	0	0	2	1	0	0	0	0

Supplement III: Comparison of oxidative pentose phosphate pathway models

The compartmentalization of the glycolysis and the oxidative pentose phosphate (PPP) pathways in plants has not been well defined. For instance, the recent models for developing sunflower embryos (Alonso *et al.*, 2007), maize root tips (Alonso *et al.*, 2007), and Arabidopsis cell culture (Williams *et al.*, 2008) consisted the oxidative and non-oxidative steps of the PPP to the plastid compartment. However, models for developing soybean cotyledons (Iyer *et al.*, 2008; Sriram *et al.*, 2004) and Arabidopsis embryos (Lonien and Schwender, 2009) included the oxidative and non-oxidative pathways of the PPP to both cytosol and plastids compartments. In recent studies of enzyme activities of Arabidopsis seed indicate evidence of the oxidative PPP operate in both cytosol and plastid compartments (Wakao *et al.*, 2008). Nonetheless, the non-oxidative PPP in the cytosol is not clearly define and it is still under investigation (Kruger and von Schaewen, 2003). Furthermore, the Rubisco (ribulose 1,5-bisphosphate carboxylase/oxygenase) bypass pathway has been proposed to contribute to carbon conversion efficiency in rapeseed and soybean (Allen *et al.*, 2009; Schwender *et al.*, 2004). Therefore, we modified our soybean metabolic network model to mimic the soybean model that was proposed by Allen et al 2009 for non-oxidative PPP with Rubisco in the plastid compartment whereas the cytosol contains both oxidative and non-oxidative PPP. Utilizing the steady-state carbon labeled metabolic flux analysis (^{13}C-MFA) of the same plant tissue, soybean zygotic embryo, (Iyer *et al.*, 2008; Sriram *et al.*, 2004) similar experiments were conducted with uniformly labeled sucrose and labeling data , and simulated fitted models in which the compartmentation of the PPP could be manipulated. The first

model is our current model with oxidative and non-oxidative in cytosol and plastid compartments acting as the base line for comparisons with the other two models. The second model is similar to the previous model with addition of Rubisco bypass reaction in the plastid. The last model is the complete PPP in cytosol, with non-oxidative and Rubisco bypass reactions were constrained in the plastid.

Simulations were performed for three models for 500 runs, each run with random start values of free flux, which was confined within the network topology (stoichiometry and carbon atom rearrangements). To evaluate the best fit of the metabolic network, a comparison of the predicted and measured isotopomer abundances for all metabolites analyzed between the three models was presented in Figure SIII-4. The best-fit flux solutions presented here were based largely on protein hydrolysate labeling data with 10% uniform labeled sucrose, hexose and pentose phosphate pools from hydrolysis of glycosylated protein and starch labeling data. The result revealed that the highest fit, based on the linear correlation coefficient, is PPP with Rubisco model ($R^2 = 0.97$), followed by PPP model ($R^2 = 0.95$), and the non-oxidative PPP with Rubisco model ($R^2 = 0.89$). Further details into which amino acids were different three models, Table SIII-1 displayed differences in relative multiplet intensities of amino acids and their corresponding precursors. Each carbon atom in an proteinogenic amino acid can be traced back to a corresponding carbon in the metabolic intermediates (Szyperski, 1995). Therefore, the table lists the metabolic intermediates of the six amino acids, which derive from the plastids, and hexose and pentose sugar pools in the cytosol and plastids compartment indicated differences in relative abundance intensity. Analysis of the hydrolysis of glycosylated protein and

starch reveal a specific compartmentalization of the soybean (Sriram *et al.*, 2007). The non-oxidative PPP with Rubisco model displayed a high degree of error between the simulation and the measurement values of hexose and pentose sugar pool. Furthermore, the chi-square error value (χ^2), which is the function that simulates isotopomer and experimental measured isotopomers, of three models were recorded. The data indicated that there is no significant difference between the χ^2 error value of PPP and PPP with Rubisco model, 106 ± 21 and 117 ± 25, respectively. The results indicate that the PPP and PPP with Rubisco models were indistinguishable. Although the flux of Rubisco bypass is nearly zero, adding the Rubisco bypass to the model reduced the variance of the majority of the flux in the network. Based on our current knowledge that Rubisco activity is present in green oil seed (Ruuska *et al.*, 2004). However, based on the proteinomic studies, there is an insignificant level of Rubsico bypass in soybean (Agrawal *et al.*, 2008). We investigate further into the isotopomer balance of the formation of Rubisco reaction as described below in Supplemental V. The non-oxidative PPP with Rubisco in the Plastid displayed the highest χ^2 value, 230 ± 427. The high variation in the χ^2 of the model indicated that the flux values for the network are unidentifiable, which could be due to the incorrect network topology. By examination of the non-oxidative PPP mechanism reactions, as suggested by Allen et al 2009, the following stoichiometries are in the plastid compartment as shown in Figure SIIIc:

$$F6P + T3P \rightarrow E4P + P5P \qquad (1)$$

$$E4P + T3P \rightarrow S7P \qquad (2)$$

$$T3P + S7P \rightarrow P5P + P5P \qquad (3)$$

Even though the carbon rearrangement in the reactions are balanced, the mechanisms were not in agreement with the stoichiometries of non-oxidative of PPP with Rubisco bypass reactions that were purposed by previous plant metabolic network model, Schwender et al 2004, which are described below:

$$F6P + T3P \rightarrow E4P + P5P \qquad (1)$$

$$F6P + E4P \rightarrow T3P + S7P \qquad (2)$$

$$T3P + S7P \rightarrow P5P + P5P \qquad (3)$$

Therefore, the high statistic value of this model could possibly be due to incorrect network topology for the soybean embryos, suggested by Allen et al, 2009. Furthermore, the isotopomer carbon-labeling pattern of the experimented and simulation model did not support the proposed network topology. Therefore, this model provides a less reliable flux network and should be rejecting based on the good fit of the available data.

Table SIII1: Metabolite(s) associated with different in experiment and simulation relative intensity

PPP model	**PPP with Rubisco model**	**Non-oxidative PPP with Rubisco model**	**Precursors**
Lys γ	Lys γ		Oxaloacetate (Plastid)
His β	His β		Ribulose-5-phospate (Plastid)
	Ile α	Ile α	Oxaloacetate (Plastid)
Pro γ	Pro γ		Glutamate (Plastid)
Asp α	Asp α	Asp α	Oxaloacetate (Plastid)
		Gxx 1	Fructose-6-phosphate (Cytosol)
		Gxx 3	Fructose-6-phosphate (Cytosol)
		Gxx 6	Glucose-6-phosphate (Cytosol)
Asp β			Oxaloacetate (Plastid)
Thr α			Oxaloacetate (Plastid)
		Starch 1	Glucose-6-Phosphate (Plastid)
		Starch 3	Glucose-6-Phosphate (Plastid)
Gxx 3			Glucose-6-Phosphate (Cytosol)

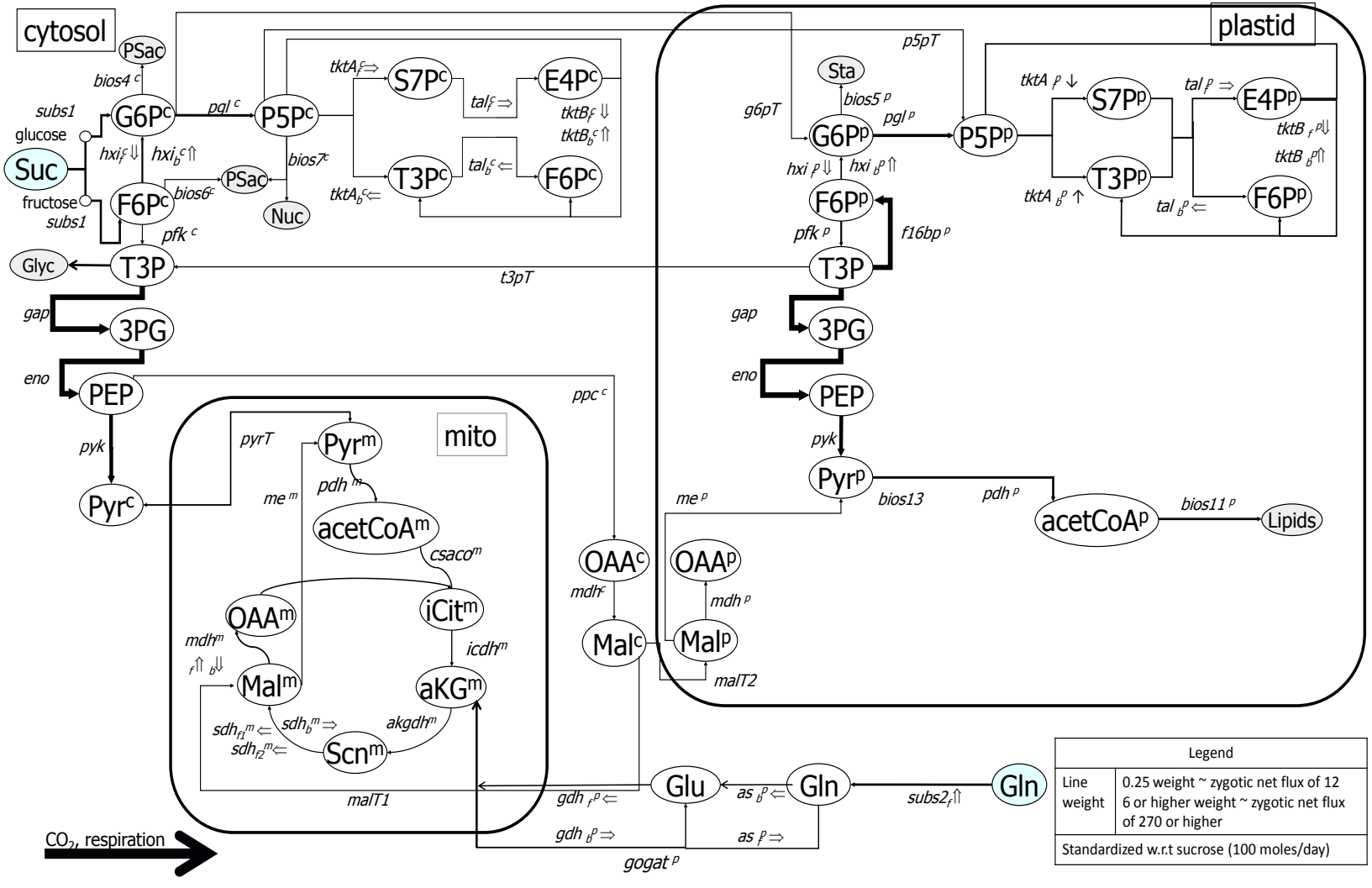


Figure SIIIa

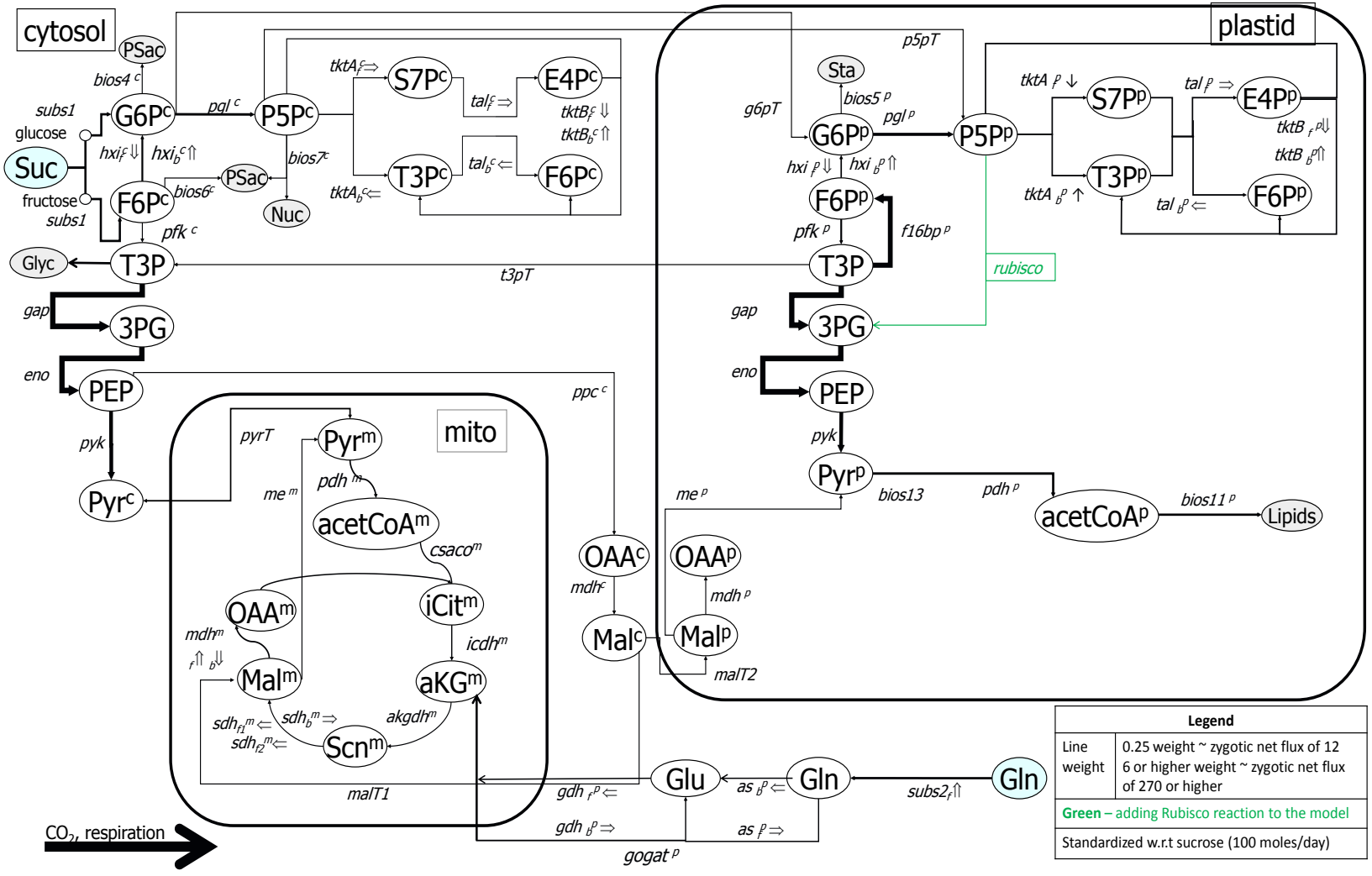


Figure SIIIb

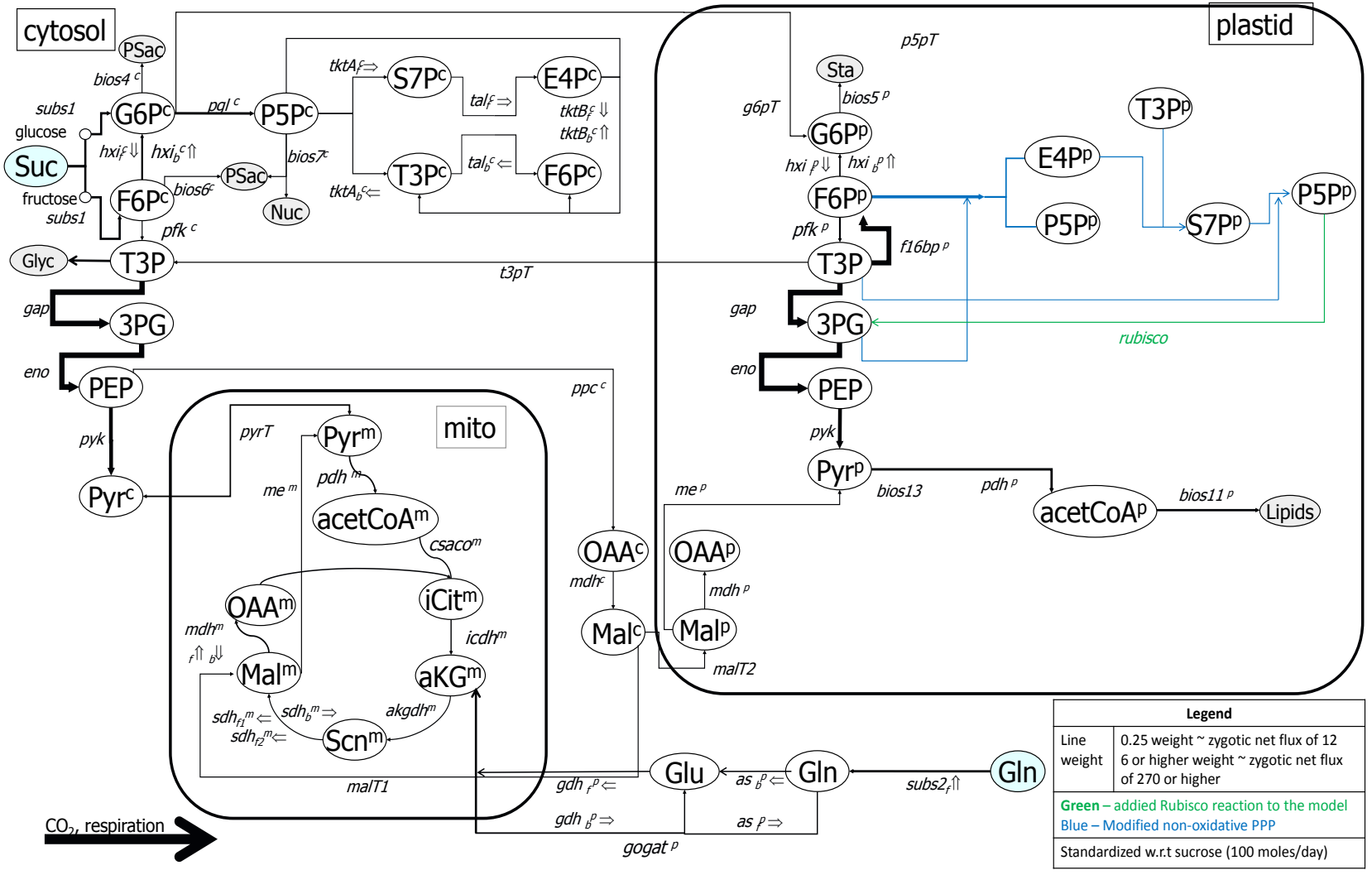


Figure SIIIc

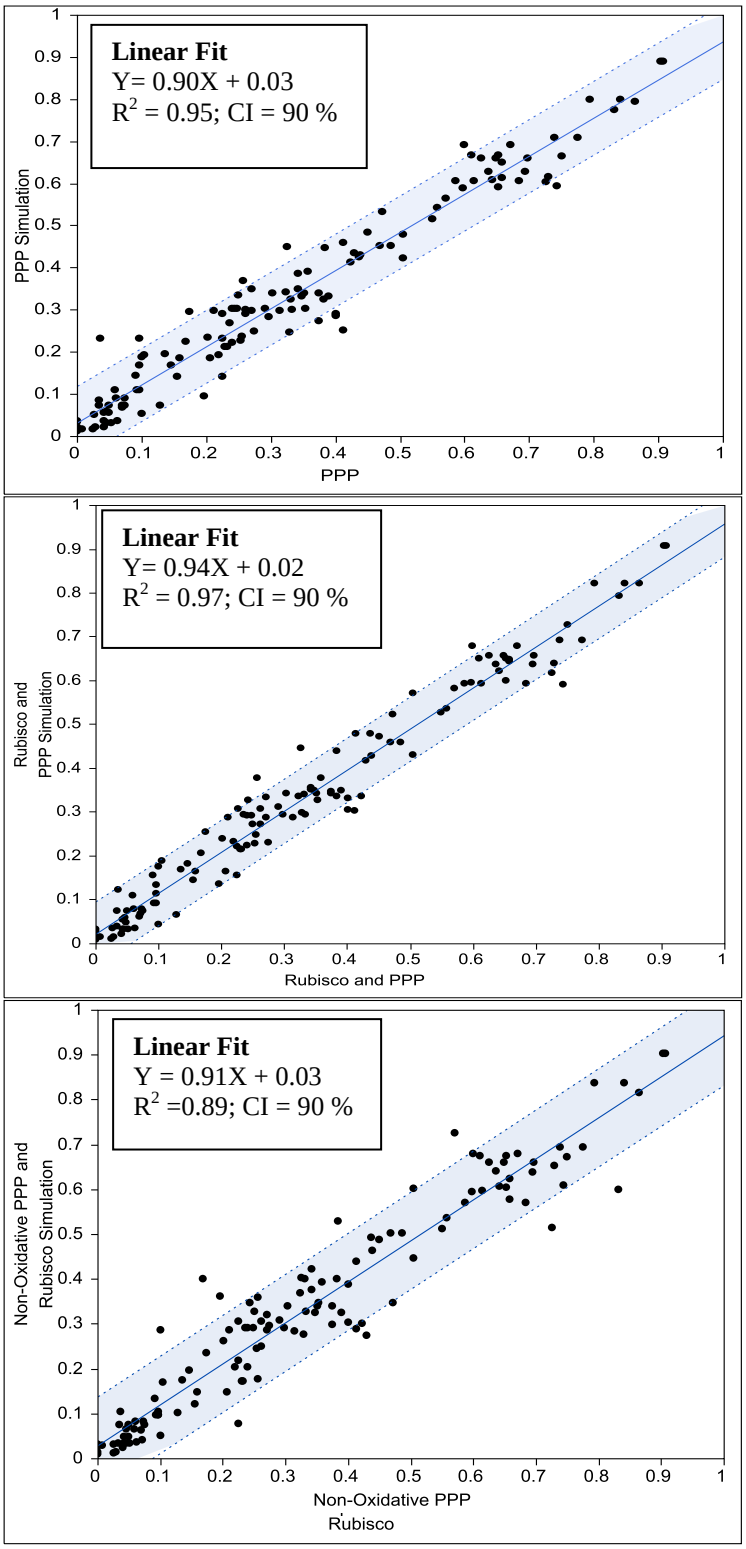


Figure SIII 4

Supplement IV: Isotopomer Balance of Rubisco reaction

The Rubisco (ribulose 1,5-bisphosphate carboxylase/oxygenase) bypass pathway has been proposed to contribute to carbon conversion efficiency in rapeseed and soybean (Allen *et al.*, 2009; Schwender *et al.*, 2004). It was not a component of the original NMR2Flux model, but was added into the network in this study to assess its contribution to flux into the storage pools. In the plastid of OxPPP, a rearrangement of carbon skeleton converts the six-carbon sugar glucose-6-phoshate (G6P) to the five-carbon sugar pentose-5-phosphate (P5P) with the formation of one CO_2; then the catalysis of the carboxylase activity of Rubisco, which fixes one molecule CO_2 with one of P5P yielding two molecules three-carbon phosphoglycerate (3PG) (Plant Metabolic Pathway Databases, htt://www.plantcyc.org). The carbon dioxide formation was in the OxPPP and fixation was in Rubisco shunt in the plastidic compartment as shown in Figure 3. To evaluate the carbon arrangement by Rubisco reaction; the overall isotopomer balance was examined. The reaction involves the formation of 3PG and the depletion of P5P in the plastid compartment as shown below:

$$\frac{d[3PG]}{dt} = \frac{d[T3P+P5P+CO_2]-PEP}{dt}$$

To write all of the isotopomer balances associated with the Rubisco reaction, we need to balance 50 isotopomers for the three-carbon position (32 from P5P, 8 from T3P, 2 from CO_2, and 8 from PEP) for the metabolic network model, including the measureable and immeasurable from the labeling patterns. In our soybean model, we have 23 measurable isotopomer balances. However, the Rubisco activity results in distribution of label from uniformly labeled sucrose into the first carbon position of

3PG. This lead to the labeling of the amino acids considered to be responsible for the Rubisco activity: phenylalanine, tryptophan, tyrosine, and serine. Examination of the first carbon position of labeling amino acid serine α (d1) indicates that the level of carbon incorporation into that isotopomer is statistically insignificant based on the student's t mean comparison analysis, 0.12 ± 0.02 and 0.14 ± 0.01, for zygotic and SHaM embryos respectively, as shown in Table S1. Since the labeling for both soybean cultures is 10% and we have natural abundance of nearly 1.1%, the total label carbon is 0.11. A similar result can be concluded with the labeling aromatic amino acids. In addition, another labeling in oxaloacetate-derived metabolites, such as threonine and related amino acids, can result from the Rubisico bypass fixation from the formation of carboxylation of PEP. Investigation of the first carbon position of labeling amino acid threonine α (d1) show that the levels were 0.10 ± 0.05 and 0.12 ± 0.06 for zygotic and SHaM embryos, respectively. Therefore, based on our isotopomer experimental design of 10% U-13C sucrose, our data indicate that we cannot identify the flux through Rubisco pathway.

Supplement V: Modify network with Transcriptomic data and measurement values

In this model, we utilized transcriptomic data of soybean zygotic and SHaM embryos (data not shown) to modify the network. Even though the the γ-Aminobutyric acid (GABA) shunt is believe to be a critical pathway for carbon and nitrogen metabolism in plant system (Fait *et al.*, 2008), transcriptomic data shows that there is evidence for glutamine assimilation via the glutamine:2-oxoglutarate amidotransferase (GOGAT) cycle and not through the GABA shunt. In addition, our metabolic flux model of uniformly labeled sucrose for soybean embryos indicates high variation in the GABA shunt pathway. This shows that the pathway is unidentifiable that could due to the limit labeling substrates that can provide greater number of isotopomers in the GABA shunt network. Recent evidence from computational design, with different combinations of the isotopomer labeling experiment studies, suggests that the best combination label experiments of 100 % 2-^{13}C Alanine and 100% U-^{13}C Glutamine, which could provide a highly identifiable GABA shunt pathway (Nargund and Sriram, 2013). Furthermore, the transcript profile data also indicated that no reserve reaction of cytosol fructose-1,6-bisphosphate adolase (pfp^c) is presented. Recent studies of transcriptomic data indicate that the glyoxylate cycle is only active during soybean seed germination or seedling development stage (Gonzalez and Vodkin, 2007) and insignificantly flux during soybean maturation development (Allen and Young, 2013). Therefore, we decided to not include the glyoxylate cycle in our metabolic network model. In summary, we removed the glyoxylate cycle, GABA shunt, and pfp^c from our network. To further simplify our network model, we eliminated the effluxes of succinate,

pyruvate and malate base upon measurement via HPLC, which indicates no present of those three components in the media after 6-day culture. Furthermore, two more reactions were added in this model: (1) the glycerol moiety (into lipid synthesis) through triose-3-phosphate (T3P) as an efflux with the constrain of 0 to 5% of total lipid extraction mimicking the flux network from soybean, Arabodopsis, and Brassica napus embryos (Allen *et al.*, 2009; Lonien and Schwender, 2009; Schwender *et al.*, 2003); (2) glycine synthesis from threonine via threonine aldolase, which converts threonine to glycine and acetaldehyde (Joshi *et al.*, 2006).

To verify the modification of our network model, we examined the statistical criteria, chi-square (χ^2) value for each manipulation of the network. Through a χ^2 fitting process, the values for fluxes (the computational variables) were optimized to recapitulate both the experimentally measured label and flux data, resulting in "best fits" of global flux values. The combination of isotopic labeling experiments with metabolic flux analysis provided a quantitative description of the flow of metabolites through biochemical pathways. Table below listed the χ^2 value for each elimination reaction from the network model and $\chi^2 = 120 \pm 35$.

Network Modification	**Model χ^2 value**
Removed succinate, pyruvate, and malate effluxes	128 ± 31
Removed glyoxylate shunt from TCA cycle	125 ± 27
Removed the reaction of cytosol fructose-1,6-bisphosphate adolase	125 ± 19
Removed the GABA shunt	110 ± 15
Added the biosynthesis of glycine from Threonine	115 ± 20
Added glycerol biosynthesis into the network	112 ± 10

Supplement VI: Detailed comparison of zygotic embryos flux map (*cv. Jack* versus *cv.* Amsoy)

Zygotic Embryos *cv Jack*,	**Allen et. al., 2009**
I. Soybean Culture	**I. Soybean Culture**
• Soybean (*Glycine max* cv. Evans)	• Soybean (*Glycine max* cv. Amsoy)
• Chamber condition: 27 ºC/20 ºC day/night temperature, 14 hour photoperiod, 500 umol photons m^{-2} sec^{-1} for light intensity	• Chamber condition: 27 ºC temperature, 15 hour photoperiod, (does not mention light intensity just stated 'sunlight supplemented by lamps as necessary')
• Seeds harvest time: 21 Day After Flower; pods were harvested from the central nodes	• Seed harvest time: R5 – 5.5 stage (~beginning seed, stage length average 15 days)
• Culture condition: 26 ºC day/night temperature, 100 umol photons m^{-2} sec^{-1} light intensity, 100 rpm	• Culture condition: 26 – 27 ºC day/night temperature, 35 umol photons m^{-2} sec^{-1} light intensity, (does not mention agitation rate)
• Cotyledons dry weight approximately 18 – 20 mg (therefore 36 – 40 mg Dwt of embryo)	• Embryo dry weight approximately 8 mg
• One cotyledon was placed in 4 ml liquid media for 6 days and the media was replaced every three days	• One cotyledon was placed in 15 ml liquid media for 14 days
• Eight cotyledons were cultured and then they were pooled after harvest to have enough material for NMR analysis.	• Three embryos (equivalent to 6 cotyledon) were cultured for each condition (biological triplicate) then embryos were pooled after harvest to have enough biomass for analysis.
II. Media components	**II. Media components**
• Sucrose – 146 mM (10% U-^{13}C label) It was indicated by Lonien and Schwender 2009, the ratio of Sucrose to glucose in the media effect the growth of the Arobidopsis embryos. Increasing Sucrose-to-Glucose ratios, embryos final size were decreased and less starch was	• Sucrose – 140 mM • Glucose – 70 mM (100% U-$^{13}C_6$ label) • Glutamine – 35 mM (100% U-$^{13}C_5$ label); Reason: Allows to test whether there is a significant flux from C4 acids into plastidic PGA, as required if scrambling via the PPP is to explain the lebeling from $^{14}CO_2$ into plastid-derived pools of PGA, PEP, or

	detected. The authors tested embryo grown in sucrose only and result indicated starchless. Furthermore, at maturation, the sucrose The result suggests that sucrose as sole sugar carbon source in culture supports embryo maturation better than glucose. The similar conclusion can be draw to soybean embryos. It has been reported that sucrose represented over 90% of all carbon in seed cat exudates, the fluid feeding embryos *in situ* (Rainbird *et al.,* 1984). Therefore, countless *in vitro* studies on soybean embryos using sucrose as a major carbon source in the media (Obendorf and Wettlaufer, 1991; Saravitz *et al.,* 1995; Pipolo *et al.,* 2004). In addition, since during soybean seed maturation sucrose becomes the dominant (Obendorf *et al.,* 1998), therefore, it was clear to use sucrose as sole sugar carbon source in the labeling experiments to obtain similar condition from *in planta* This explained why the embryos from Allen et al 2009 paper were growing slow than our culture. The soybean embryos in this experiment uptake equal amount of glucose as compared to sucrose even	pyruvate (resulting in labeling of serine, fatty acids, and aromatic amino acids) Results: (1) There was no significant presence of multiple labeled molecules in these metabolite pools when embryos were labeled with glutamine, supporting the idea that Rubisco and the Calvin cycle rather than gluconeogenesis and PPP are responsible for the results of $^{14}CO_2$ experiment (described below); (2) The hexose-derived products (starch and cell wall) show little or no detectable labeling, indicating a lock of gluconeogenic activity in expanding soybean embryos in light. • Asparagine – 12.5 mM **Radio-labeled Carbon Balance Study** Cultures were pre-incubated in unlabeled medium for 5 days and then transferred aspetically to equivalently radiolabeled medium; then the cultures were maintained in this environment for 5 days. • ^{14}C-Labeled Carbon sources: [U-$^{14}C_{12}$]-Sucrose (495 mCi mmol^{-1}), [U-$^{14}C_6$]-glucose (245 mCi mmol^{-1}), [U-$^{14}C_5$]-glutamine (219 mCi mmol^{-1})

though the concentrations of sucrose was much higher than glucose in the media (140 mM sucrose versus 70 mM glucose). This indicates that embryos preferred transferred the monosaccharide carbon source through the membrane rather than disaccharide, which requires less energy during the transport process. Furthermore, the experiment had 100% uniformly labeled glucose and 100% nonlabeled sucrose in the media. The flux map indicates that both carbon sources called "sugars" as a input into the network. ^{13}C-MFA tracks labeled carbon in the pathway, in this case labeled glucose, but we have sucrose in the media as well and embryo did use sucrose as a carbon source. How can be differentiated which carbon that was consumed by soybean embryos? How are the regulations and pathways different with sucrose and/or glucose in plant carbon metabolism? (Could be different co-factors, ATP, NADPH depend on the carbon source) • Glutamine – 37 mM **III. Detection methods** • 2D [^{13}C, ^{1}H] HSQC spectroscopy for protein and starch hydrolysates labeled • HPLC for Sucrose and	**III. Detection methods** • Combination NMR and GC-MS spectroscopy for protein, lipids, carbohydrates and organic acids • NMR spectroscopy for starch hydorlysate

Glutamine uptake (in house measurement and from DuPont as well) • Oil – gravimetrically using hexane extraction and NMR analysis *in situ* • Total protein content – CHN Elementary analysis • Proteinogenic amino acid – HPLC (data were from Purdue University) • Starch content quantification – starch assay kit **IV. Calculation uptakes** • Sucrose and glutamine - determined from before and after culture • Sucrose and glutamine – determine by taking sample daily (with a different set of nonlabeled experiment) • Both calculations were compared and it was no significantly different. **V. Simulation program** • Fluxes were quantified – generic flux evaluation software NMR2Flux (Sriram *et al*., 2004) **VI. Metabolic network/map** • LVA peaks from 2 separate metabolic flux analyses (protein and starch hydrolysate) from the soybean embryos can distinguish between glycolysis and PPP pathway in two compartments (cytosol and plastid) • The RuBisCo bypass is located in the plastid. From extracellular flux data to estimate CO_2 fixation by RuBisCo is quite small therefore did not include in this work	• 1D [^{1}H] NMR – for sucrose, glucose, glutamine and asparagines uptake • Oil – quantified as FAMEs using GC-FID • Total protein content – CHN Elementary analysis **IV. Calculation uptakes** • Carbon and nitrogen - determined from before and after culture **V. Simulation program** • Fluxes were quantified – 13CFLUX suite of software provided by W. Wiecher (Wiechert *et al*., 2001) **VI. Metabolic network/map** • Labeling plastidic (starch) and cytosolic (cell wall and proteoglycan) carbohydrate pools was found to be similar therefore cannot be distinguish between two compartments • Role of Rubisco in the metabolism of soybean was indicated by $^{14}CO_2$ and ^{13}C labeling data – oxidative pentose phosphate pathway is a major flux • 11% of the CO_2 released by lipid

• It is unlikely the CO_2 from the cytosol/mitochondrion could have been fixed by RuBisCo because this CO_2 has a labeling is substantially lesser than 11% (since glutamine contributes to it), and if it were fixed we should observe that the labeling of Leu δ^2 and Val γ^2 were lesser than 0.12 (was not observed in our case) • Metabolic Flux Map: Figure 1 • Raw Data from absolute and relative fluxes are listed in Table 1 **VII. Biomass composition Results 10% U-13C sucrose** • Growth (mg Dwt/cotyledon-day): 8.78 ± 1.74 Day 0 = 17.13 ± 2.45 Dwt mg/cotyledon; Day 6 = 69.80 ± 11.57 mg Dwt/cotyledon Relative growth rate: 0.52 ± 0.10 Dwt per day • Protein Dwt (%): 37.81 ± 0.06 • Lipid Dwt (%): 15.22 ± 0.13 • Starch Dwt (%): 9.88 ± 0.47 • Soluble sugar Dwt (%): 6.29 ± 0.68 • Ash + DNA/RNA Dwt (%): 13 • Residual Biomass Dwt (%):18.20 ± 2.12 A set of nonlabeled sucrose experiment was performed and the results of the growth rate, sucrose & glutamine uptake rates, and biomass composition were equivalent to the labeled. Therefore, the labeled sucrose did not have effect on the macro levels of the biomass composition.	synthesis and the TCA cycle – from $^{14}CO_2$ labeling patterns supported role of RuBisCo-based • Metabolic Flux Map: Figure 2 • Flux data indicated in the map **VII. Biomass composition results Culture [U-13C] glutamine and [U-13C] glucose study** • Growth (mg Dwt/embryo-day): 6.8 ± 0.7 Day 0 = 8 mg Dwt; Day 14 = 103 mg Dwt Relative growth rate: 0.43 ± 0.04 Dwt per day • Protein (%): 39.3 ± 3.6 (glucose + glutamine experiments) (GCMS); • Lipid (%): 18.5 ± 0.9 (glucose + glutamine experiments) (GCMS) **Notice:** ***The relatively intensity of both experiments (glucose and glutamine) are dramatic different, Table SIc) and lot of data are missing or undetectable for protein and lipid labeling profiles*** • Starch (%): not provided; • Cell Wall, Protein Glycan, and Starch labeling (glucose experiment) (NMR) ***Notice: Relatively intensity of Hexose metabolites were detected in glucose experiment only and again the data are not completed, Table SId)*** • Sucrose and glucose uptake

• Sucrose uptake (umol/cotyledon-day): 68.5 ± 7.2 Day 0 = 146 mM; Day 3 = 122 mM; Day 6 = 120 mM [Replace media after three day, even after 3 days we always have ~ 80 % of the sucrose concentration in the media] • (68*6day) = 0.408 mmol sucrose/cotyledon • Initial media: (146*8) = 1.168 mmol sucrose • To calculate yield = (74 – 17 mg Dwt)/(0.408 mmol sucrose) = 140 mg Dwt/mmol sucrose • Glutamine uptake (umol/cotyledon-day): 29.75 ± 1.2 (29*6day) = 0.174 mmol glutamine/cotyledon To calculate yield = (74 – 17 mg Dwt)/(0.174 mmol glutamine) = 328 mg Dwt/mmol glutamine Day 0 = 37 mM ; Day 3 = 22 mM; Day 6 = 18 mM	(umol/embryo-day): 17.0 ± 1.3 and 16.9 ± 3.0, respectively [Therefore, approximately 68 umol sugars/cotyledon-day] (68*14day) = 0.952 mmol sugars/embryo • Initial media: (140*15ml) = 2.1 mmol sucrose and (70*15 ml) = 1.05 mmol glucose. Total = 3.15 mmol • Sucrose uptake = (34*14) = 0.476 mmol; Glucose uptake = (34*14) = 0.476 mmol • Sucrose in the media after 14 days assume no volume change = (2.1 – 0.476) = 1.624 mmol = (1.624/0.015) = 108 mM • Sucrose: Day 0 = 140 mM and Day 14 = 108 mM • Glucose in the media after 14 days assume no volume change = (1.05 – 0.476) = 0.574 mmol = (0.574/0.015) = 38 mM • Glucose: Day 0 = 70 mM and Day 14 = 38 mM • To calculate yield = (103 – 8 mg Dwt)/(0.952 mmol sugars) = 99.8 mg Dwt/mmol sugars • Glutamine uptake (umol/embryo-day): 14.3 ± 1.0 [Therefore it is ~ 28 umol/Cotyledon-day] (28*14day) = 0.392 mmol glutamine/embryo Initial media: (35*15) = 0.525 mmol glutamine Glutamine in the media after 14 days assume no volume change = (0.525-0.392) = 0.133 mmol = (0.133/0.015) = 8.87 mM • To calculate yield = (103 – 8 mg Dwt)/(0.392 mmol sucrose) = 242 mg Dwt/mmol glutamine • Glutamine: Day 0 = 35 mM and Day 14 = 8.87 mM

	Notice: After 14 day culture the embryos utilized nearly 50% of glucose and 75% glutamine

Supplement VII: Extended procedure of proteinogenic amino acid and starch labeling of 2D NMR

In this section, we describe an ideal protocol of extraction of the biomass components of proteinogenic amino acid and starch labeling from soybean embryo culture. Generating sufficient NMR samples can have a high cost due to the labeling substrates. Moreover, for each NMR sample, approximately 20-50 mg of material is required to obtain an adequately resolved 2D HSQC NMR spectrum with a sufficient signal-to-noise ratio for quantification. Therefore, proper techniques and careful planning for NMR sample preparation are presented. Furthermore, the section provides summary of the 2D NMR analysis, which includes the techniques used to obtain the relative intensities of the 16 amino acids and glucosyl units measured from protein fraction, as well as the determination of the latter in the starch fraction of soybean embryos.

1. Materials

1. Freeze dry system (Labconco, Kansas City, MO; cat. no. 77500-00)

2. Rapidvap evaporator (Labconco, Kansas City, MO; cat. no. 7900002)

3. Hydrolysis tube 20-ml (Pierce, Rockford, IL; cat. no. 29564)

4. Hydrochloric acid (constant boiling) 6N Sequencing Grade (Pierce, Rockford, IL; cat. no. 24308)

5. Deuterium oxide, D_2O, (Sigma-Aldrich, St. Louis, MO; cat. no. CAS 7789-20-0)

6. Deuterium chloride, DCl, (Sigma-Aldrich, St. Louis, MO; cat. no. CAS 7698-05-7)

7. Dialysis Cassette (2,000 MWCO) (Pierce, Rockford, IL; cat. no. 66212)

8. Heating block (Pierce, Rockford, IL; cat. no. 18870)

9. 3-(Trimethylsilyl)-1-propanesulfonic acid sodium salt (TMS) (Sigma-Aldrich, St. Louis, MO; cat. no. CAS 2039-96-5)

10. NMR tubes (Kimble-Chase, Vineland, NJ; Kontes Article no. 897240-0000)

11. Coomassie Plus (Bradford) Protein Assay (Pierce, Rockford, IL; cat. no. 23236)

12. β-Mercaptoethanol (Biorad, Hercules, CA; cat. no. 1610710)

13. Amyloglucosidase from *Aspergillus niger* (Sigma-Aldrich, St. Louis, MO; cat. no. CAS 9032-08-0)

14. Glucose Assay Kit (Sigma-Aldrich, St. Louis, MO; cat. no. GAGO20)

15. Soda Lime (VWR, Radnor, PA; cat. no. CAS 8006-28-8)

16. 50-ml plastic tubes with screw caps (Falcon™ Conical Centrifuge Tubes, BD Biosciences, San Jose, CA)

17. Vacuum Pump (model RV5 115/230V, 1-ph, 50/60Hz) (Edwards, Crawley Sussex, England)

18. -80°C Freezer (So-Low Environmental Equipment, Cincinnati, OH)

19. Spin-X® Centrifuge Tube Filter (0.22 µm nylon, Costar, Corning, NY; cat. no. 8169)

20. Refrigerated Microcentrifuge Model 5415R (Eppendorf)

21. Precision Water Bath Model 25 (Thermo Scientific, cat. no. 51221080)

22. Water bath sonicator (Fisher Scientific FS 110H Ultrasonic Cleaner)

23. Geno/Grinder® (SPEX SamplePrep, Metuchen NJ, USA)

2. Methods

The powerful NMR technique does not require any pretreatment method, such as derivatization, prior to analysis, which makes standardization straightforward.

However, NMR samples have to have a sufficient amount of materials for the analysis of intracellular fluxes, which are the extraction of protein and starch fractions for 2D NMR analysis. One challenge in the extraction protocol is to have sufficient amount of materials for NMR analysis. Our extraction method is unique because oil, protein, soluble sugars, and starch can sequentially in a separation process scheme, be extracted from one biomass sample (Figure 1).

The biomass fractionation of soybean embryos provides a way to extract individual soluble metabolites separately. This prevents the complication of mixing crude extracts from the tissue, which could potentially interfere with the analysis of carbon coupling patterns in 2D HSQC NMR spectra. Furthermore, the proteinogenic amino acids observed in the spectra of protein hydrolysates illustrates a unique method that is able to distinguish the intracellular compartmentation within the tissue. One of many examples, the biosynthesis pathway of the amino acid histidine is located inside the plastid (Stepansky and Leustek, 2006); therefore, histidine plays an important role in compartment identification between plastid and cytosol. Another unique aspect of our method is that the glucosyl units observed in the 2D HSQC NMR spectra of protein and of starch hydrolysate directly reflect glucose-6-phosphate pools in the cytosol and the plastid compartments within the cell, respectively. As mentioned in Sriram et al. (10), acid hydrolysate of soybean protein contains hydrolysis products of glucose and mannose. Starch is a glucose polymer; therefore, acid hydrolysis of soybean starch yields glucose monomers.

The analysis of the carbon coupling patterns from the signals of 2D HSQC NMR spectra of the proteinogenic amino acids and starch materials from the extract

requires the highest available purity. Therefore, consider autoclaving or sterilizing all of the glass bottles that will contain buffer and reagents materials before proceeding with the protocol.

2.1 Biomass extraction, Hydrolysis, Amino Acid Quantification

1. Soybean *in vitro* cultures are harvested by pouring the contents of each culture into a sieve, rinsing with 100 ml distilled water, blotting dry, and weighing and freezing in liquid nitrogen prior to lyophilization at -50°C and 0.0158 mbar for 72 hours. The lyophilized embryos are finely grounded using a Geno/Grinder® prior to further analysis.

2. Approximately 100 mg of dry powdered soybean embryo sample is extracted with 1 ml of n-hexane at 40°C for 1 hour; the extracts are centrifuged at high speed (20,000 g) for 10 minutes at room temperature. The process is repeated five times and after each extraction, the solvent containing the lipids is pooled into a pre-weighed glass tube (approximately 4.8 ml final volume) and dried for four days in a hood at room temperature. The mass of lipids after solvent evaporation is measured gravimetrically.

3. The remaining hexane-extracted biomass is dried and further extracted for protein in 600 µl of 200 mM phosphate buffer (pH = 7.2) containing 14 mM β-mercaptoethanol at 4°C for 20 minutes to dissolve and suspend the protein in solution. Then the solution is centrifuged at high speed (13,000 rpm) at 4°C for 15 minutes and the supernatant is transferred into a micro-centrifuge tube. The extraction is performed on the pellet two more times with 400 µl of the buffer (per extraction) and the supernatants are pooled with that from the initial extraction

(approximately 1.2 ml final volume). Protein contents of the extract are measured by Bradford Assay (Bradford, 1976)*see* **Note 1**).

4. Residual soluble sugars in the defatted/deproteinated-biomass pellets are extracted into 1 ml of 80% aqueous ethanol in a water bath sonicator at 60 °C for 20 minutes, vortexed every 5 minutes. The extraction is repeated four times and after each extraction, the solvent containing the soluble sugars are pooled into a pre-weighed glass tube and dried for 3 days in a 40 °C oven, and the soluble sugar content in these fractions is measured gravimetrically.

5. The remaining pellets of each sample, after soluble sugars extraction in step 4, are placed in 1.5 ml of Distilled De-Ionized (DDI) water, covered with aluminum foil, and autoclaved (liquid cycle at 121 °C and a pressure of 15 psi) for 30 minutes prior to starch digestion and extraction (*see* **Note 2**).

6. Starch is digested in 1.5 ml of 100 mM citrate buffer (pH 5.0) containing amyloglucosidase at a ratio of 0.025 mg enzyme/ 1 mg of tissue dry weight (approximately 2 µl of amyloglucosidase or 0.6 units of enzyme are added to each sample). Samples are incubated in a 30°C water bath overnight. The extractions from step 5 are centrifuged at high speed for 30 minutes at room temperature. Starch content is quantified using a glucose assay kit (*see* **Materials**).

7. Extracted starch sample is frozen in a -80 °C freezer for 4 hours and lyophilized at -50 °C and 0.0158 mbar for 72 hours (*see* **Note 3**). Then, starch powder is dissolved in DDI water.

8. Protein extract is dialyzed using dialysis cassettes in 10 mM phosphate buffer (pH 7.2) at 4 °C for 24 hours to remove any soluble sugar residues in the sample (*see* **Note 4**).

9. The protein and starch samples are vacuum hydrolyzed in hydrolysis tubes using 6N constant boiling hydrochloric acid (HCl). The acid is added in the ratio of 6 ml HCl to 10 mg protein/starch.

10. The hydrolysis tubes of protein and starch samples are evacuated multiple times purged with nitrogen to remove residual air, and re-evacuated to prevent oxidation during hydrolysis.

11. The protein and starch samples in the hydrolysis tubes are hydrolyzed for 4 hours at 140-145°C in a heat block (*see* **Note 5**).

12. After hydrolysis, a Rapidvap evaporator is used to evaporate the hydrochloric acid for 2 hours at 35 % speed, 40 bar pressure, and a temperature of 45°C (*see* **Note 6**).

13. Each dried hydrolysate (one for protein and one for starch) is reconstituted in 2 ml of de-ionized water using a Rapidvap evaporator for 30 minutes at 70 % speed, atmospheric pressure, and room temperature (*see* **Note 7**).

14. The reconstituted samples are filtered using centrifuge and Spin-X® centrifuge tube with filter (*see* **Note 8**). The filters are discarded, the protein and starch samples tubes are tightly sealed with parafilm, and holes are poked for vaporization of water during freeze-drying.

15. The protein and starch samples are frozen in a -80°C freezer for 4 hours and lyophilized at -50°C and 0.0158 mbar for 72 hours or until the samples are completely dried.

16. The protein and starch samples are reconstituted in 400 µl of deuterium oxide and vortexed. Then, a small portion of protein sample is diluted to 20-50 µl using DDI water and amino acids are analyzed by HPLC. 16 Amino Acids can be analyzed by using pre-column derivatization with 6-Aminoquinolyl-N-Hydroxysuccinimidyl carbamate (Cohen, 2000).

2.2 NMR Sample Preparation (Protein and Starch)

Protein and starch samples are reconstituted with deuterium oxide as described in the previous section. Next, adjust the pH of the sample to 0.8 – 1.0 using DCl/D_2O mixture if required; this step is crucial due to chemical shifts of the carbon and hydrogen is highly depending on pH of the samples. A pinch of TMS standard flakes are added, vortexed well, and then transferred to clean NMR tubes.

2.3 2-D HSQC Spectral Analysis of Amino Acids

1. 2-D [^{1}H, ^{13}C] Heteronuclear Single Quantum Correlation (HSQC) spectroscopy of hydrolyzed protein and starch samples are performed on a Bruker Avance DRX 500 MHz spectrometer at 298 K. The HSQC spectra were acquired using a modified version of the INEPT spectra (insensitive nuclei enhanced by polarization transfer) pulse sequence (Bodenhausen and Ruben, 1980). TMS is used as an internal standard: the reference zero p.p.m. is set by the methyl signal of TMS. The magnetic resonance frequency of ^{13}C (F1 dimension) and ^{1}H (F2 dimension) are 125.7 MHz and 499 MHz, respectively. The spectral widths along the F1 and F2 dimensions are 5028.05 Hz and 5482.26 Hz, respectively. Peak aliasing is applied to minimize the sweep width in the F1 dimension. The number of complex data points is 900 (^{13}C) and 1024 (^{1}H). The number of scans ranges from 16 to 32.

2. NMR spectra are acquired using Xwinnmr (Bruker) software and analyzed by using the free software NMRview (Johnson and Blevins, 1994). NMRview is available in both Linux and Window versions.

3. Cross peaks of s, d, and t can be processed directly to obtain the raw signal intensities by using the NMRview software.

4. Both aliphatic and aromatic carbon atoms of 16 amino acids in the spectra are identified from hydrolysate of the soybean protein samples, as shown in Figures 2 and 3, respectively (Sriram *et al.*, 2004).

5. Each carbon of the amino acid listed in the figures is detectable by its unique ^{13}C and ^{1}H chemical shifts (Wuthrich, 1976) as well as distinctive coupling patterns and J-coupling constants (Jcc) (Harris, 1983). The chemical shifts and the J-coupling-constant (Jcc) values obtained from (Krivdin and Kalabin, 1989) are organized in Table 1 (*see* **Note 9**).

6. Table 1 displays the relationship of the ^{13}C labeling pattern in the multiplet to carbon atoms. A multiplet can contain one or any combination of the following: a singlet “s”, a doublet “d”, a doublet split by a smaller scalar coupling “d1”, a doublet split by a larger scalar coupling “d2”, a doublet of doublets “dd”, and a triplet “t” (Szyperski, 1995).The metabolic precursor to the given metabolite measured are included in Table 1 as well.

7. Furthermore, Table 1 includes the multiplet component of the cross-peak to the specific isotopomer. For example, singlet “s” under “Cross-Peak” corresponds to the first isotopomer listed under “Isotopomer”. The notation for each of these isotopomers indicates the labeling state of the observed carbon “Carbon number” and

its neighbor(s). The boldfaced carbon atom(s) in Isotopomer indicates labeled carbon atom(s), normal font indicates unlabeled carbon atom(s), and x indicates an unknown labeling state of the carbon atom(s).

8. To quantify overlapping multiplets [(s, d1, d2, dd), (s, d1, d2, d3, dd1, dd2, dd3, qd)] on the HSQC spectra, peak deconvolution software was developed in our lab, based on a spectral model originally proposed by van Winder et al. (36).

9. Since the peaks of the aromatic amino acids and some of the glucosyl units are crowded, in order to separate the resonances additional 2D HSQC spectra are acquired that are J-scaled along the ^{13}C dimension, by integrating pulse sequences reported previously (Brown, 1984; Willker *et al.*, 1997). J-scaling increases multiplet separation by an even integral factor J and eliminates multiplet overlap. J-scaling factors of six are employed in processing aromatic amino acid and some of glycosyl units and will be described below (Sriram *et al.*, 2007).

10. The multiplet intensities (s, d, t) provided by the NMRView program are raw intensities. However, NMR2Flux (11), the simulation program used to calculate the metabolic fluxes, requires intensity fractions.

11. Table 2 shows the equation that converts the raw intensities of the carbon atom of the amino acids and their noise data to the relatively intensities of the carbon atoms amino acids and noise (*see* **Note 10**)

2.4. 2-D HSQC Spectra Analysis of Glucosyl Units

1. Figure 2 displays the 2D HSQC spectrum of 16 amino acids and glucosyl units, hydroxyacetone (HyA) and levulinic acid (LVA), from soybean protein hydrolysate (*see* **Note 11**).

2. The reaction products of the hydrolysis of the sugar residues in the glycoprotein are HyA and LVA (Sriram *et al.*, 2007).

3. Table 3 shows glucosyl units, carbon number, cross-peak, precursors, and isotopomers of the protein and starch hydrolysates of soybean embryo culture.

4. The HyA1 cross peak multiplet (s, d) of both protein and starch hydrolysates can be quantified by NMRview software on 2D HSQC spectrum to obtain raw intensities.

5. To quantify the overlapping multiplets (LVA 3, LVA 4, LVA 6 of protein and starch hydrolysates) in the HSQC spectra, a peak deconvolution software is written based on a spectral model originally proposed by van Winden et al. (van Winden *et al.*, 2001).

3. Notes

1. The total soluble protein and starch contents of the cultures are estimated at 50 mg each, which should be more than sufficient for NMR analysis. As a result, a total of ~ 300 mg of dry powdered soybean embryo will need to be processed in multiple extractions of 100 mg each, and the content pooled for analysis, rather than scaling up the extraction process for one-step. As expected, the Bradford Assay result from extractable protein is lower than the method of determining the total protein content of the embryos via elemental analysis. Therefore, it is crucial to measure the amount of protein that can be extracted from the tissue.

2. Besides autoclaving starch samples, it is helpful to have positive/negative starch control samples. Prepare positive starch control samples, simply by placing 10-15 mg waxy corn starch in 3 – 12x75 mm glass tubes (weigh tube accurately), add 1.5 ml DDI water to each of the three tubes, vortex samples until dissolved, cover with aluminum foil. For negative starch control samples, just place 1.5 ml DDI water to each of the 3-12x75 mm glass tubes and cover with aluminum foil.

3. The purpose to lyophilize starch sample is to concentration and reduce the volume prior to acid hydrolyze treatment.

4. Dialysis of the protein extract is necessary to eliminate sugar molecules that may have been co-extracted with the protein. The dialysis cassette model that was listed in the material section could hold up between 3 to 12 ml in volume; however, during the injection process of sample to the cassette, air can easily be trapped in the cassette. Therefore, a maximum of 6 ml in volume per cassette should be placed.

Place the 6 ml-protein-cassette into the 2 liters of 10 mM phosphate buffer with pH 7.2 at 4 °C; replace the buffer every 12 hours.

5. Turn on the heating block sometime in advance so that the temperature is 140-145°C before hydrolyzing your protein and starch samples. Also, before closing the hydrolysis tube make sure that no acid sticks to the top of the tubes; it can damage the Teflon in the cap. A maximum volume of the sample to be stored in hydrolysis tube is 6 ml ($1/3^{rd}$ of the tube size) since the vacuum and heat cause the level to rise, which possibility causes the content to pull up.

6. After 4 hours of incubation the protein and starch samples are removed from the heating block and let at room temperature. Then, place samples on ice or in the refrigerator before opening caps. Be extremely careful before opening the cap since the tube samples are under considerable pressure. Soda lime can be used to trap acid from the protein and starch acid hydrolyzed samples. If the Rapidvap step is not being done on the same day, store the hydrolysis tube in the refrigerator without releasing the vacuum.

7. After the protein and starch samples have been placed in the high temperature Rapidvap evaporator, these liquid samples are nearly evaporated. Therefore, add DDI water to reconstitute sample thoroughly.

8. The maximum volume capacity of the Spin-X® micro-centrifuge tube with filter is less than 0.8 ml due to the insertion of the filter. Therefore, make sure not to add more than 0.8 ml of protein/starch acid hydrolysis to the filter insertion of the micro-centrifuge tube.

9. Table 1 lists each carbon of the 16 amino acids; however, some of the carbon atoms of the amino acid provide a similar information regarding the precursor of the metabolic pathway. Therefore, the extra measurements of the amino acid can be used to check the validity of that particular flux in the metabolic pathway.

10. Taking a noise measurement is necessary for each peak which has its intensity quantified. For this purpose, the intensity of about 10 noise peaks that are closed to the target peak is averaged. Please verify that the sum of the fractional intensities of s, d, and t add up to 1. The raw intensity of the cross-peak is then input in the Excel spread sheet form.

11. For clarity, hydrolyzed protein extract contains HyA and LVA peaks (located in the cytosol compartment) and hydrolyzed starch extract also contains HyA and LVA peaks (located in the plastid compartment).

Figure Captions

Figure 1: Schematic diagram of biomass fractionation of soybean dry tissue and preparation of protein and starch hydrolysates for NMR analysis.

Figure 2: Detailed scheme for analyzing the carbon coupling patterns in 2D [^{1}H, ^{13}C] HSQC spectrum of 16 proteinogenic amino acids and glucosyl units of soybean (*Glycine max)* zygotic embryo *cv.* Jack.

Figure 3: 2D HSQC spectrum of aromatic rings of protein extracted from soybean zygotic embryo *cv. Jack*.

Figure 4: 2D HSQC spectrum of hydrolyzed starch samples from soybean zygotic embryo *cv*. Jack.

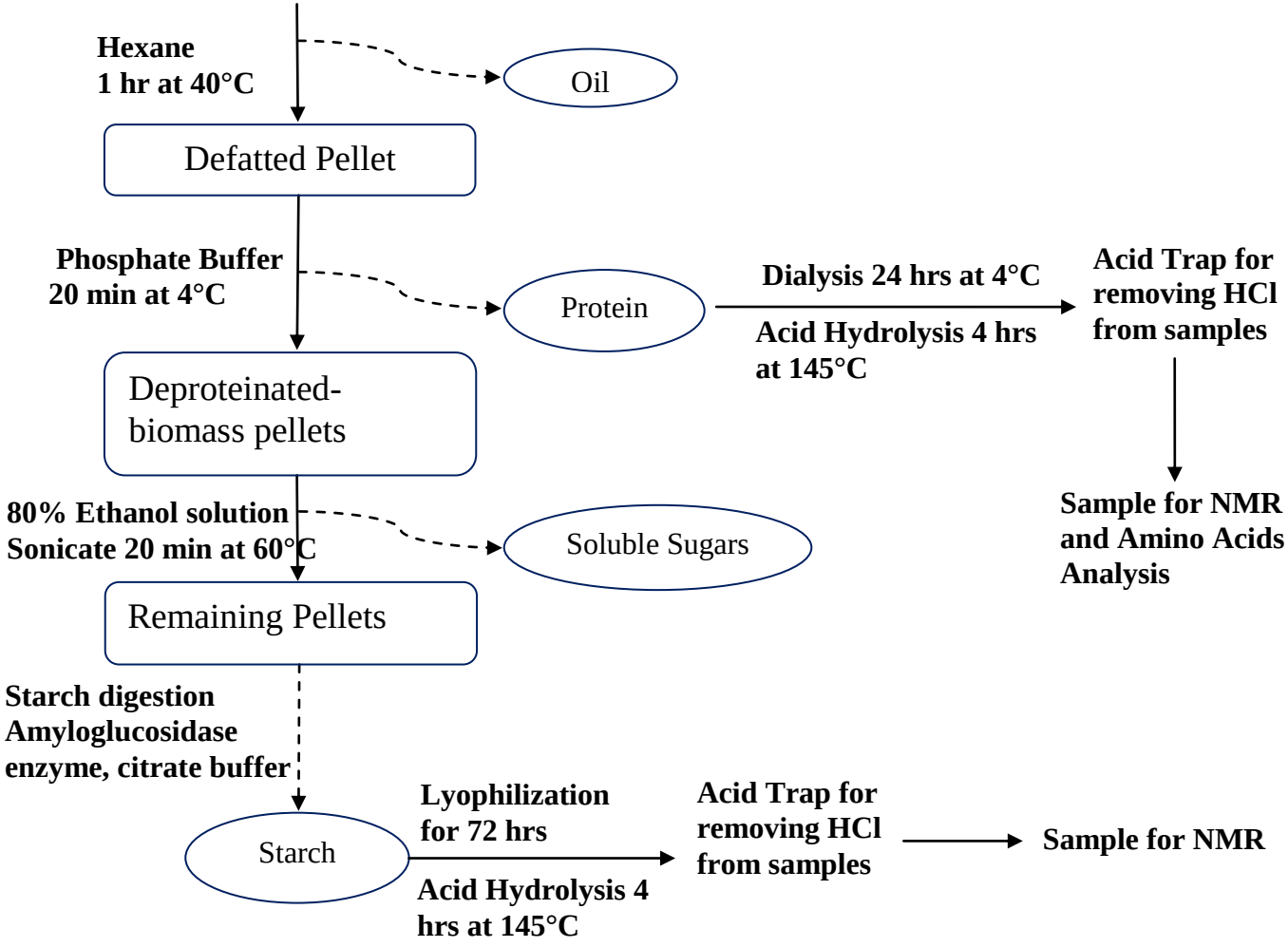

Hexane
1 hr at 40°C
Oil
Defatted Pellet
Phosphate Buffer
20 min at 4°C
Protein
Dialysis 24 hrs at 4°C
Acid Hydrolysis 4 hrs
at 145°C
Acid Trap for
removing HCl
from samples
Sample for NMR
and Amino Acids
Analysis
Deproteinated-
biomass pellets
80% Ethanol solution
Sonicate 20 min at 60°C
Soluble Sugars
Remaining Pellets
Starch digestion
Amyloglucosidase
enzyme, citrate buffer
Starch
Lyophilization
for 72 hrs
Acid Hydrolysis 4
hrs at 145°C
Acid Trap for
removing HCl
from samples
Sample for NMR

Figure 1:

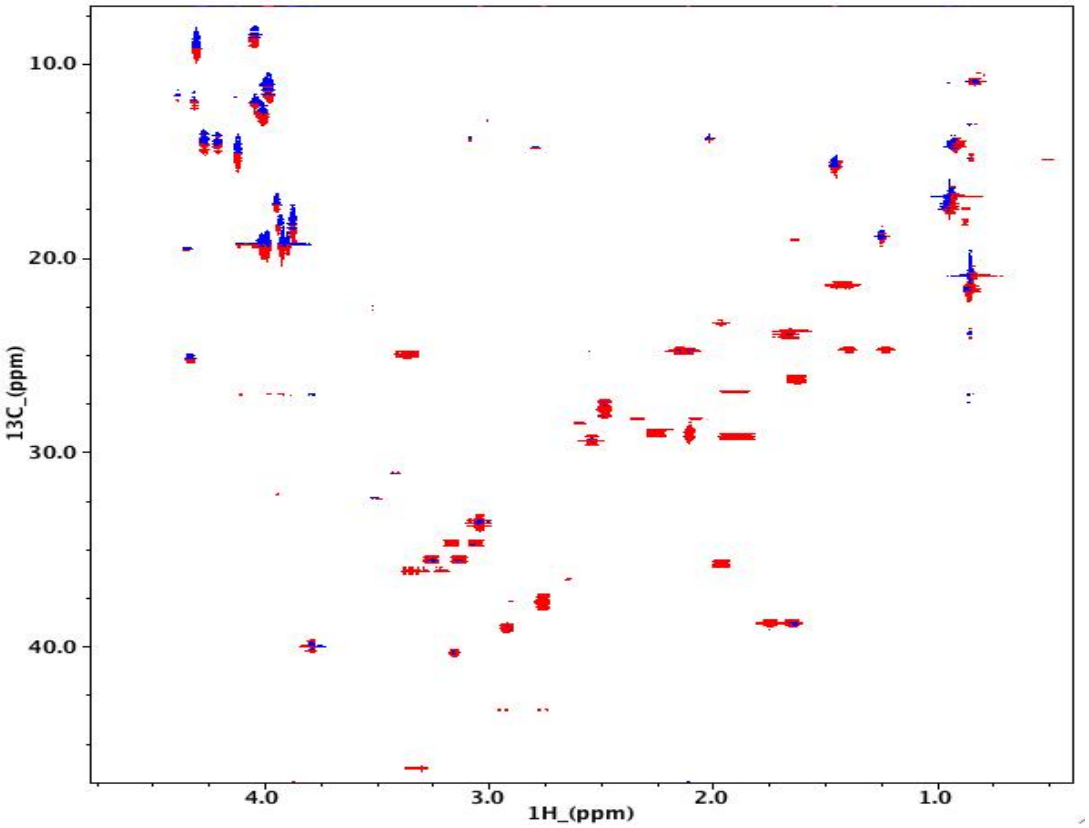


Figure 2

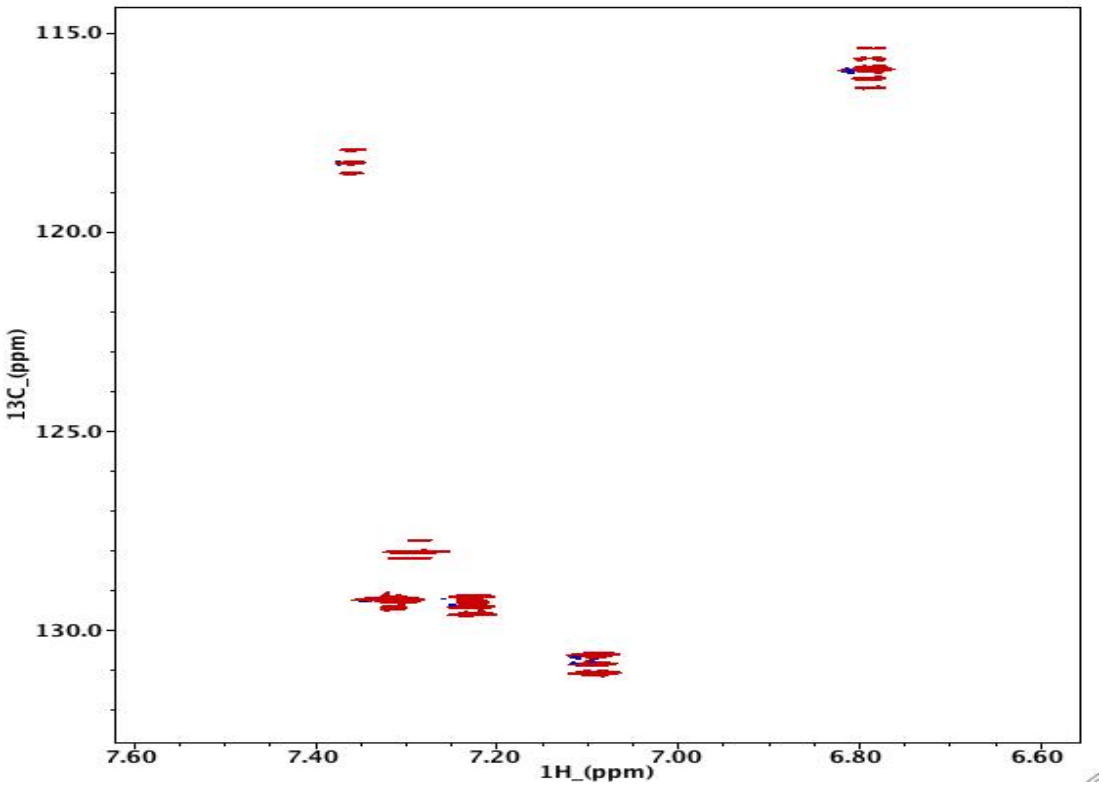


Figure 3

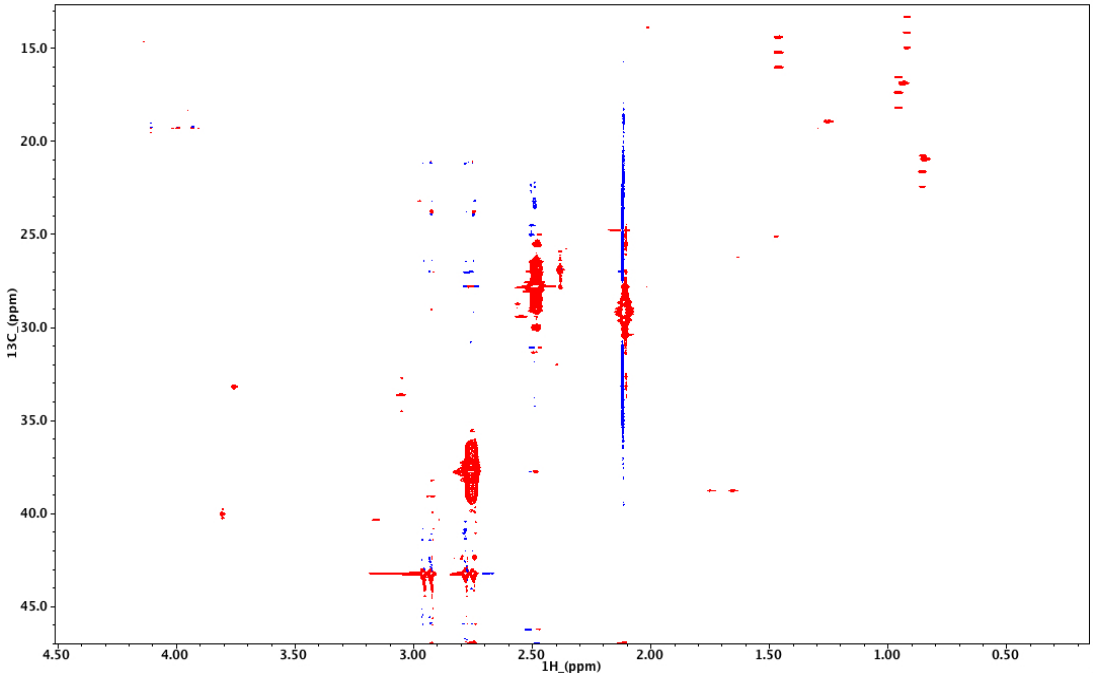

15.0
20.0
25.0
30.0
35.0
40.0
45.0
13C_(ppm)
4.50
4.00
3.50
3.00
2.50
2.00
1.50
1.00
0.50
1H_(ppm)

Figure 4

Table Captions

Table 1. The cross-peaks to carbon atoms, chemical shift values for the 16 amino acids, and their J-coupling constant from a 2D HSQC NMR spectrum of soybean embryos. The notation of the cross-peak multiplet: s indicates singlet, d indicates doublet, d1 and d2 indicate the first and second doublet, dd indicates double doublet, and t indicates triplet. The precursor corresponding to each amino acid is indicated and its subcellular localization is referred in parentheses. The boldfaced carbon atom(s) in isotopomer indicates the labeled carbon atom(s), normal font indicates unlabeled carbon atom(s), and x indicates unknown labeling state of the carbon atom(s).

Table 2. Equations to determine each fraction intensity from multiplet s, d, t, and noise fraction from 2D HSQC NMR spectrum.

Table 3. Carbon number, cross-peak multiplet, precursors, and Isotopomers of glucosyl units of protein and starch hydrolysate samples from soybean culture.

Table 1:

Amino Acid	Carbon Position	Carbon Number	Cross-Peaks	1-H	13-C	Jcc (d1, d2)	Precursor	Isopotomer
Ala	α	2	s, d1, d2, dd	4.08, 4.12	11.54	59.6, 34.3	Pyruvate	[1**2**3], [**12**3], [1**23**], [**123**]
	β	3	s, d	1.59	17.98	34.3		[x2**3**], [x**2**3]
Arg	α	2	s, d1, d2, dd	4.18	14.69	53.4, 33.7	Glutamate	
	β	3	s, d, t	1.97	29.69	33.7, 34.1		[x2**3**4x], [x**23**4x]+[x2**34**x], [x**234**x]
	γ	4	s, d, t	1.67	26.52	34.1, 35.8		
	δ	5	s, d	3.23	43.07	35.8		[xxx4**5**], [xxx**45**]
Asp	α	2	s, d1, d2, dd	4.34	12.15	59.8, 37.6	Oxaloacetate	[1**2**3x], [**12**3x], [1**23**x], [**123**x]
	β	3	s, d1, d2, dd	3.1	36.5	37.6, 55.5		[x2**3**4], [x**23**4], [x2**34**], [x**234**]
Glu	α	2	s, d1, d2, dd	4.49	14.94	59.6, 33.7	Glutamate	
	β	3	s, d, t	2.21	27.6	33.7, 36.3		[x2**3**4x], [x**23**4x]+[x2**34**x], [**x234x**]
	γ	4	s, d1, d2, dd	2.61	32.18	36.3, 54.7		[xx3**4**5], [xx**34**5], [xx3**45**], [xx**345**]
Gly	α	2	s, d	3.85	42.84	59.6	3-phosphoglycerate	[1**2**x], [**12**x]
His	α	2	s, d1, d2, dd	4.4	14.93	59.8, 34.6	Ribose-5-phosphate (Plastid)	[xx3**4**5], [xx**34**5], [xx3**45**], [**xx345**]
	β	3	s, d1, d2, dd	3.42	27.86	51, 34.6		[x2**3**4x], [x**23**4x], [x2**34**x], [x**234**x]
	δ2	4	s, d	7.55	40.87	74.5		[**1**2xxx], [**12**xxx]

Table 1: (continued.)

Ile	α	2	s, d1, d2, dd	3.91	21.18	59.2, 32.5	Pyruvate (Plastid)	[1**2**xx]·[x2x], [**12**xx]·[x2x], [1**2**xx]·[x**2**x], [**12**xx]·[x**2**x]
	β	3	s, d1, d2, dd	2.04	38.52	32.5, 33.9		
	γ1	4	s, d, t	1.26, 1.46	27.46	33, 34.9		[x2x]·[xx**3**4], [x**2**x]·[xx**3**4] + [x2x]·[xx**34**], [x**2**x]·[xx**34**]
	γ2	5	s, d	1	16.92	33.9		[x2**3**], [x**23**]
	δ	6	s, d	0.92	13.67	34.9		[xx3**4**], [xx**34**]
Leu	α	2	s, d1, d2, dd	4.04	14.29	59.6, 32.8	Pyruvate (Plastid)	[1**2**]·[x2x], [**12**]·[x2x], [1**2**]·[x**2**x], [**12**]·[x**2**x]
	β	3	s, d, t	1.72, 1.83	41.63	32.8, 32.4		[x2]·[x**2**x]·[x2x], [x**2**]·[x**2**x]·[x2x] + [x2]·[x2x]·[x2x], [x**2**]·[x**2**x]·[x**2**x]
	γ	4	s, d1, d2, dd	1.75	26.72	32.4, 34.3		
	δ1	5	s, d	0.94	24.36	34.3		[x2**3**], [x**23**]
	δ2	6	s, d	0.93	23.65	34.3		[x2x]·[xx**3**], [x**2**x]·[xx**3**]
Lys	α	2	s, d1, d2, dd	4.05	15.54	53.6, 34.2	Oxaloaceate/Pyruvate	
	β	3	s, d, t	1.93, 1.99	32.02	34.2, 34.5		½{[x2**3**4] + [x2**3**]·[xxx4]}c, ½{[x**23**4]+[x2**34**] + [x**23**]·[xxx4] +

Table 1: (continued.)

								[x2**3**]·[xxx**4**]}, ½{[x**234**]+[x**23**]·[xxx**4**]
	γ	4	s, d, t	1.48	24.16	34.5, 34		[xx3**4**]·[xx3], [xx**34**]·[xx3] + [xx3**4**]·[xx**3**], [xx**34**]·[xx**3**]
	δ	5	s, d, t	1.69	29.03	34, 35.4		½{[x2**3**4] + [x2**3**]·[xxx4]}, ½{[x**23**4]+[x2**34**] + [x**23**]·[xxx4] + [x2**3**]·[xxx**4**]}, ½{[x**234**] + [x**23**]·[xxx**4**]}
	ε	6	s, d	2.99	41.75	35.4		½{[x**2**3]+[x**2**3x]}, ½{[x**23**]+[x**23**x]}
Met	α	2	s, d1, d2, dd	3.97	21.28		Oxaloaceate (Plastid)	
	β	3	s, d1, d2, dd	2.25	31.63			
	γ	4	s, d	2.6	31.29			[xx3**4**], [xx**34**]
	ε	5	s, d	2.1	16.62			
Phe	α	2	s, d1, d2, dd	4.32	16.93	59.9, 32.3	Phosphoenolpyruvate	[1**2**3], [**12**3], [1**23**], [**123**]
	β	3	s, d1, d2, dd	3.34, 3.21	38.32	32.3, 43.7		[x2**3**]·[x2x], [x**23**]·[x2x], [x2**3**]·[x**2**x], [x**23**]·[x**2**x]
	γ	4	s, d, t	6.99, 7.36	25.08	43.7		
	ε	5	s, d, t	6.66,	20.03			

Table 1: (continued.)

				7.07				
Pro	α	2	s, d1, d2, dd	4.39	22.53		Glutamate (Plastid)	
	β	3	s, d, t	2.15, 2.41	31.07			
	γ	4	s, d, t	2.03	26.12			[xx3**4**5], [xx**34**5]+[xx3**45**], [xx**345**]
	δ	5	s, d	3.39	48.97			[xxx4**5**], [xxx**45**]
Ser	α	2	s, d1, d2, dd	4.31	15.05	59.2, 36.4	Serine	[1**2**3], [**12**3], [1**23**], [**123**]
	β	3	s, d	3.98, 4.06	22.1	36.4		[x2**3**], [x**23**]
Thr	α	2	s, d1, d2, dd	4.17	17.56	58.8, 36.3	Oxaloaceate (Plastid)	[1**2**3x], [**12**3x], [1**23**x], [**123**x]
	β	3	s, d, t	4.34, 4.39, 4.44	28	36.6, 37.5		
	γ2	4	s, d	1.32	21.67	37.5		[xx3**4**], [xx**34**]
Tyr	α	2	s, d1, d2, dd	4.26	17.03	53.2, 33	Phosphoenolpyruvate	[1**2**3], [**12**3], [1**23**], [**123**]
	β	3	s, d, t	3.25, 3.13	37.5	33, 33.7		[x2**3**]·[x2x], [x**23**]·[x2x] + [x23]·[x2x], [x**23**]·[x**2**x]
	δ1	5	s, d, t	6.99, 7.34	13.6	57.5, 58		[x23]·[1xxx], [x**23**]·[1xxx]+[x2**3**]·[**1**xxx], [x**23**]·[**1**xxx]
	δ2	5	s, d, t	6.99,	13.6	57.5, 58		[x2x]·[**4**3xx],

Table 1: (continued.)

				7.34				[x**2**x]·[**4**3xx]+[x2x]·[**43**xx], [x**2**x]·[**43**xx]
	ε1	6	s, d, t	6.77, 6.85, 6.96	38.69	58, 60.5		[xx3]·[**1**2xx], [xx**3**]·[**1**2xx] + [xx3]·[**12**xx], [xx**3**]·[**12**xx]
Val	α	2	s, d1, d2, dd	4	20.14	59.1, 32.2	Pyruvate (Plastid)	[1**2**x]·[x2x], [**12**x]·[x2x], [1**2**x]·[x**2**x], [**12**x]·[x**2**x]
	β	3	s, d1, d2, dd	2.25	31.77	34.4, 33.7		
	γ1	4	s, d	1.04	20.14	34.4		[x23], [x**23**]
	γ2	5	s, d	1.02	19.56	33.7		[x2x]·[xx**3**], [x**2**x]·[xx**3**]

Table 2

Multiplet Type	**Singlet**	**Doublet**	**Triplet**
s, d	$\frac{s\ intensity}{total\ intensity}$	$\frac{d\ intensity + d\ intensity}{total\ intensity}$	
s, d, t	$\frac{s\ intensity - (t\ intensity + t\ intensity)}{total\ intensity}$	$\frac{d\ intensity + d\ intensity}{total\ intensity}$	$\frac{2 * (t\ intensity + t\ intensity)}{total\ intensity}$
Noise fraction	$\frac{n\ estimated\ intensity}{total\ intensity}$		

Table 3

Hydrolysate	Glucosyl Unit	Carbon Number	Cross-peak	Precursors	Isopotomer
Protein	HyA 1	6	s, d	Glucose-6-phosphate/Fructose-6-phosphate (Cytosol)	[**1**2xxxx], [**12**xxxx]
Protein	LVA 3	6	s, d1, d2, dd	Glucose-6-phosphate/Fructose-6-phosphate (Cytosol)	[x2**3**4xx], [x**23**4xx], [x2**34**xx], [x**234**xx]
Protein	LVA 4	6	s, d1, d2, d3, dd1, dd2, dd3, qd	Glucose-6-phosphate/Fructose-6-phosphate (Cytosol)	[xx3**4**56], [xx**34**56], [xx3**45**6], [xx3**45**6], [xx**345**6], [xx**34**5**6**], [xx3**456**], [xx**3456**]
Protein	LVA 6	6	s, d1,d2, dd	Glucose-6- phosphate /Fructose-6-phosphate (Cytosol)	[xxx45**6**], [xxx4**56**], [xxx**45**6], [xxx**456**]
Starch	HyA 1	6	s, d	Glucose-6-phosphate (Plastid)	[**1**2xxxx], [**12**xxxx]
Starch	LVA 3	6	s, d1, d2, dd	Glucose-6-phosphate (Plastid)	[x2**3**4xx], [x**23**4xx], [x2**34**xx], [x**234**xx]
Starch	LVA 4	6	s, d1, d2, d3, dd1, dd2, dd3, qd	Glucose-6-phosphate (Plastid)	[xx3**4**56], [xx**34**56], [xx3**45**6], [xx3**45**6], [xx**345**6], [xx**34**5**6**], [xx3**456**], [xx**3456**]
Starch	LVA 6	6	s, d1,d2, dd	Glucose-6-phosphate (Plastid)	[xxx45**6**], [xxx4**56**], [xxx**45**6], [xxx**456**]

References

Agrawal GK, Hajduch M, Graham K, Thelen JJ. 2008. In-depth investigation of the soybean seed-filling proteome and comparison with a parallel study of rapeseed. *Plant Physiology* 148, 504-518.

Allen DK, Libourel IGL, Shachar-Hill Y. 2009. Metabolic flux analysis in plants: coping with complexity. *Plant Cell and Environment* 32, 1241-1257.

Allen DK, Ohlrogge JB, Shachar-Hill Y. 2009. The role of light in soybean seed filling metabolism. *Plant Journal* 58, 220-234.

Allen DK, Shachar-Hill Y, Ohlrogge JB. 2007. Compartment-specific labeling information in C-13 metabolic flux analysis of plants. *Phytochemistry* 68, 2197-2210.

Allen DK, Young JD. 2013. Carbon and Nitrogen Provisions Alter the Metabolic Flux in Developing Soybean Embryos. *Plant Physiology* 161, 1458-1475.

Allen DK, Young JD. 2013. Carbon and Nitrogen Provisions Alter the Metabolic Flux in Developing Soybean Embryos. *Plant Physiology*.

Allen SM, Butler KH, Carlson TJ, Hitz WD, Stoop JM. 2008. Plastidic phosphoglucomutase genes. Vol. US 20080066204A1: EI du Pont de Nemours and Company.

Alonso AP, Dale VL, Shachar-Hill Y. 2010. Understanding fatty acid synthesis in developing maize embryos using metabolic flux analysis. *Metabolic Engineering* 12, 488-497.

Alonso AP, Goffman FD, Ohlrogge JB, Shachar-Hill Y. 2007. Carbon conversion efficiency and central metabolic fluxes in developing sunflower (Helianthus annuus L.) embryos. *Plant Journal* 52, 296-308.

Alonso AP, Piasecki RJ, Wang Y, LaClair RW, Shachar-Hill Y. 2010. Quantifying the Labeling and the Levels of Plant Cell Wall Precursors Using Ion Chromatography Tandem Mass Spectrometry. *Plant Physiology* 153, 915-924.

Alonso AP, Raymond P, Hernould M, et al. 2007. A metabolic flux analysis to study the role of sucrose synthase in the regulation of the carbon partitioning in central metabolism in maize root tips. *Metabolic Engineering* 9, 419-432.

Alonso AP, Raymond P, Rolin D, Dieuaide-Noubhani M. 2007. Substrate cycles in the central metabolism of maize root tips under hypoxia. *Phytochemistry* 68, 2222-2231.

Alonso AP, Val DL, Shachar-Hill Y. 2011. Central metabolic fluxes in the endosperm of developing maize seeds and their implications for metabolic engineering. *Metabolic Engineering* 13, 96-107.

Bailey MA, Boerma HR, Parrott WA. 1993. Genotype-Specific Optimization of Plant-Regeneration from Somatic Embryos of Soybean. *Plant Science* 93, 117-120.

Bodenhausen G, Ruben DJ. 1980. Natural abundance nitrogen-15 NMR by enhanced heteronuclear spectroscopy. *Chemical Physics Letters* 69, 185-189.

Boghigian BA, Seth G, Kiss R, Pfeifer BA. 2010. Metabolic flux analysis and pharmaceutical production. *Metabolic Engineering* 12, 81-95.

Borisjuk L, Nguyen TH, Neuberger T, et al. 2005. Gradients of lipid storage, photosynthesis and plastid differentiation in developing soybean seeds. *New Phytologist* 167, 761-776.

Bowsher CG, Tobin AK. 2001. Compartmentation of metabolism within mitochondria and plastids. *Journal of Experimental Botany* 52, 513-527.

Boyle NR, Morgan JA. 2009. Flux balance analysis of primary metabolism in Chlamydomonas reinhardtii. *BMC Systems Biology* 3, (7 January 2009).

Bradford MM. 1976. Rapid and Sensitive Method for Quantitation of Microgram Quantities of Protein Utilizing Principle of Protein-Dye Binding. *Analytical Biochemistry* 72, 248-254.

Brown LR. 1984. Differential Scaling along Omega-1 in COSY Experiments *Journal of Magnetic Resonance* 57, 513-518.

Carrari F, Urbanczyk-Wochniak E, Willmitzer L, Fernie AR. 2003. Engineering central metabolism in crop species: learning the system. *Metabolic Engineering* 5, 191-200.

Chanprame S, Kuo TM, Widholm JM. 1998. Soluble Carbohydrate Content of Soybean [Glycine max (L.) Merr.] Somatic and Zygotic Embryos during Development. *In Vitro Cellular & Developmental Biology. Plant* 34, 64-68.

Clemente TE, Cahoon EB. 2009. Soybean Oil: Genetic Approaches for Modification of Functionality and Total Content. *Plant Physiology* 151, 1030-1040.

Cohen SA. 2000. Amino Acid Analysis Using Precolumn Derivatization with 6-Aminoquinolyl-N-Hydroxysuccinimidyl Carbamate. *Methods in Molecular Biology: Amino Acid Analysis Protocols* 159, 39-47.

Dietz K. 1987. Control function of hexosemonophosphate isomerase and phosphoglucomutase in starch synthesis in leaves. *In J Biggins, ed, Proceedings of the 7th International Congress on Photosynthesis* 3, 329-332.

Dieuaide-Noubhani M, Alonso AP, Rolin D, Eisenreich W, Raymond P. 2007. *Metabolic flux analysis: recent advances in carbon metabolism in plants*.

Durot M, Bourguignon PY, Schachter V. 2009. Genome-scale models of bacterial metabolism: reconstruction and applications. *FEMS Microbiology Reviews* 33, 164-190.

Fait A, Fromm H, Walter D, Galili G, Fernie AR. 2008. Highway or byway: the metabolic role of the GABA shunt in plants. *Trends in Plant Science* 13, 14-19.

Feist AM, Herrgard MJ, Thiele I, Reed JL, Palsson BO. 2009. Reconstruction of biochemical networks in microorganisms. *Nature Reviews Microbiology* 7, 129-143.

Finer JJ, McMullen MD. 1991. Transformation of Soybean Via Particle Bombardment of Embryogenic Suspension Culture Tissue. *In Vitro Cellular and Developmental Biology Plant* 27P, 175-182.

Gifford RM, Thorne JH. 1985. Sucrose Concentration at the Apoplastic Interface Between Seed Coat and Cotyledons of Developing Soybean Seeds. *Plant Physiology* 77, 863-868.

Golombek S, Rolletschek H, Wobus U, Weber H. 2001. Control of storage protein accumulation during legume seed development. *Journal of Plant Physiology* 158, 457-464.

Gonzalez DO, Vodkin LO. 2007. Specific elements of the glyoxylate pathway play a significant role in the functional transition of the soybean cotyledon during seedling development. *Bmc Genomics* 8.

Haga KI, Sodek L. 1987. Utilization of Nitrogen-Sources by Immature Soybean Cotyledons in Culture. *Annals of Botany* 59, 597-601.

Harris RK. 1983. Nuclear Magnetic Resonance Spectroscopy: A Physiochemical View. *Pitman Books, London*.

Hartwig EE, Kilen TC. 1991. Yield and Composition of Soybean Seed from Parents with Different Protein, Similar Yield. *Crop Science* 31, 290-292.

Hattenbach A, Heineke D. 1999. On the role of chloroplastic phosphoglucomutase in the regulation of starch turnover. *Planta* 207, 527-532.

Hay J, Schwender J. 2011. Computational analysis of storage synthesis in developing Brassica napus L. (oilseed rape) embryos: flux variability analysis in relation to (13)C metabolic flux analysis. *Plant Journal* 67, 513-525.

Hay J, Schwender J. 2011. Metabolic network reconstruction and flux variability analysis of storage synthesis in developing oilseed rape (Brassica napus L.) embryos. *Plant Journal* 67, 526-541.

Hayati R, Egli DB, CraftsBrandner SJ. 1996. Independence of nitrogen supply and seed growth in soybean: Studies using an in vitro culture system. *Journal of Experimental Botany* 47, 33-40.

He Y, Young T, Clark K, Kleppinger-Sparace K, Bridges W, Sparace S. 2011. Developmental profile of storage reserve accumulation in soybean somatic embryos. *In Vitro Cellular & Developmental Biology - Plant* 47, 725-733.

Hernandez-Sebastia C, Marsolais F, Saravitz C, Israel D, Dewey RE, Huber SC. 2005. Free amino acid profiles suggest a possible role for asparagine in the control of storage-product accumulation in developing seeds of low- and high-protein soybean lines. *Journal of Experimental Botany* 56, 1951-1963.

Hsu FC, Bennett AB, Spanswick RM. 1984. Concentrations of Sucrose and Nitrogenous Compounds in the Apoplast of Developing Soybean Seed Coats and Embryos *Plant Physiology* 75, 181-186.

Hymowitz T, Collins FI, Walker WM, Panczner J. 1972. Relationship between content of oil, protein, and sugar in soybean seed. *Agronomy Journal* 64, 613-616.

Iyer VV, Sriram G, Fulton DB, Zhou R, Westgate ME, Shanks JV. 2008. Metabolic flux maps comparing the effect of temperature on protein and oil biosynthesis in developing soybean cotyledons. *Plant Cell and Environment* 31, 506-517.

Iyer VV, Sriram G, Shanks JV. 2007. Metabolic flux maps of central carbon metabolism in plant systems - Isotope labeling analysis. *Concepts in Plant Metabolomics*, 125-144.

Johnson BA, Blevins RA. 1994. NMR View - A Computure-Program for the Visualization and Analysis of NMR data. *Journal of Biomolecular Nmr* 4, 603-614.

Joshi V, Laubengayer KM, Schauer N, Fernie AR, Jander G. 2006. Two Arabidopsis threonine aldolases are nonredundant and compete with threonine deaminase for a common substrate pool. *Plant Cell* 18, 3564-3575.

Junker BH, Lonien J, Heady LE, Rogers A, Schwender J. 2007. Parallel determination of enzyme activities and in vivo fluxes in Brassica napus embryos grown on organic or inorganic nitrogen source. *Phytochemistry* 68, 2232-2242.

Kell DB. 2004. Metabolomics and systems biology: making sense of the soup. *Current Opinion in Microbiology* 7, 296-307.

Kinney AJ. 1996. Development of genetically engineered soybean oils for food applications. *Journal of Food Lipids* 3, 273-292.

Koch K. 2004. Sucrose metabolism: regulatory mechanisms and pivotal roles in sugar sensing and plant development. *Current Opinion in Plant Biology* 7, 235-246.

Koch KE. 1996. Carbohydrate-modulated gene expression in plants. *Annual Review of Plant Physiology and Plant Molecular Biology* 47, 509-540.

Krivdin LB, Kalabin GA. 1989. Structural Applications of One-Bond Carbon-Carbon Spin-Spin Coupling-Constants *Progress in Nuclear Magnetic Resonance Spectroscopy* 21, 293-448.

Kruger NJ, Masakapalli SK, Ratcliffe RG. 2012. Strategies for investigating the plant metabolic network with steady-state metabolic flux analysis: lessons from an Arabidopsis cell culture and other systems. *Journal of Experimental Botany* 63, 2309-2323.

Kruger NJ, Ratcliffe RG. 2009. Insights into plant metabolic networks from steady-state metabolic flux analysis. *Biochimie* 91, 697-702.

Kruger NJ, Troncoso-Ponce MA, Ratcliffe RG. 2008. 1H NMR metabolite fingerprinting and metabolomic analysis of perchloric acid extracts from plant tissues. *Nat. Protocols* 3, 1001-1012.

Kruger NJ, von Schaewen A. 2003. The oxidative pentose phosphate pathway: structure and organisation. *Current Opinion in Plant Biology* 6, 236-246.

Last RL, Jones AD, Shachar-Hill Y. 2007. Towards the plant metabolome and beyond. *Nature Reviews Molecular Cell Biology* 8, 167-174.

Li ZS, Moon BP, Xing AQ, Liu ZB, McCardell RP, Damude HG, Falco SC. 2010. Stacking Multiple Transgenes at a Selected Genomic Site via Repeated Recombinase-Mediated DNA Cassette Exchanges. *Plant Physiology* 154, 622-631.

Libourel IGL, Shachar-Hill Y. 2008. Metabolic flux analysis in plants: from intelligent design to rational engineering. *Annual Review of Plant Biology* 59, 625-650.

Litterer LA, Plaisance KL, Schnurr JA, Storey KK, Jung HJG, Gronwald JW, Somers DA. 2006. Biosynthesis of UDP-glucuronic acid in developing soybean embryos: possible role of UDP-sugar pyrophosphorylase. *Physiologia Plantarum* 128, 200-211.

Lonien J, Schwender J. 2009. Analysis of Metabolic Flux Phenotypes for Two Arabidopsis Mutants with Severe Impairment in Seed Storage Lipid Synthesis. *Plant Physiology* 151, 1617-1634.

Lu C, Napier JA, Clemente TE, Cahoon EB. 2011. New frontiers in oilseed biotechnology: meeting the global demand for vegetable oils for food, feed, biofuel, and industrial applications. *Current Opinion in Biotechnology* 22, 252-259.

Manjunath S, Lee CHK, VanWinkle P, Bailey-Serres J. 1998. Molecular and biochemical characterization of cytosolic phosphoglucomutase in maize - Expression

during development and in response to oxygen deprivation. *Plant Physiology* 117, 997-1006.

Masakapalli SK, Lay PL, Huddleston JE, Pollock NL, Kruger NJ, Ratcliffe RG. 2009. Subcellular flux analysis of central metabolism in a heterotrophic Arabidopsis thaliana cell suspension using steady-state stable isotope labeling. *Plant Physiology*, pp.109.151316.

Masakapalli SK, Le Lay P, Huddleston JE, Pollock NL, Kruger NJ, Ratcliffe RG. 2010. Subcellular Flux Analysis of Central Metabolism in a Heterotrophic Arabidopsis Cell Suspension Using Steady-State Stable Isotope Labeling. *Plant Physiology* 152, 602-619.

Massou S, Nicolas C, Letisse F, Portais JC. 2007. NMR-based fluxomics: Quantitative 2D NMR methods for isotopomers analysis. *Phytochemistry* 68, 2330-2340.

Medini D, Serruto D, Parkhill J, Relman DA, Donati C, Moxon R, Falkow S, Rappuoli R. 2008. Microbiology in the post-genomic era. *Nature Reviews Microbiology* 6, 419-430.

Meyer K, Hitz WD, Yadav NS, Damude HG. 2012. DGAT genes from Yarrowia lipolytica for increased seed storage lipid production and altered fatty acid profiles in soybean. E I du Pont de Nemours and Company. US Patent No US 8143473.

Meyer K, Stecca KL, Ewell-Hicks K, Allen SM, Everard JD. 2012. Oil and Protein Accumulation in Developing Seeds Is Influenced by the Expression of a Cytosolic Pyrophosphatase in Arabidopsis. *Plant Physiology* 159, 1221-1234.

Miflin BJ, Lea PJ. 1977. Amino- Acid Metabolism *Briggs, Winslow R. (Ed.). Annual Review of Plant Physiology, Vol. 28. X+615p. Illus. Annual Reviews Inc.: Palo Alto, Calif., USA. Isbn 0-8243-0628-7*, 299-329.

Miflin BJ, Lea PJ. 1977. Amino Acid Metabolism. *Annual Review of Plant Physiology* 28, 299-329.

Miranda M, Borisjuk L, Tewes A, Heim U, Sauer N, Wobus U, Weber H. 2001. Amino acid permeases in developing seeds of Vicia faba L.: expression precedes storage protein synthesis and is regulated by amino acid supply. *Plant Journal* 28, 61-71.

Morgan JA, Barney CS, Penn AH, Shanks JV. 2000. Effects of buffered media upon growth and alkaloid production of Catharanthus roseus hairy roots. *Applied Microbiology and Biotechnology* 53, 262-265.

Nargund S, Sriram G. 2013. Designer labels for plant metabolism: statistical design of isotope labeling experiments for improved quantification of flux in complex plant metabolic networks. *Molecular Biosystems* 9, 99-112.

Nielsen J, Oliver S. 2005. The next wave in metabolome analysis. *Trends in Biotechnology* 23, 544-546.

Nishizawa K, Ishimoto M. 2009. Maturation of somatic embryos as a model for soybean seed development. *Plant Biotechnology* 26, 543-550.

Nishizawa K, Takagi K, Teraishi M, Kita A, Ishimoto M. 2010. Application of somatic embryos to rapid and reliable analysis of soybean seed components by RNA interference-mediated gene silencing. *Plant Biotechnology* 27, 409-420.

Nishizawaa K, Ishimoto M. 2009. Maturation of somatic embryos as a model for soybean seed development. *Plant Biotechnology* 26, 543-550.

Norton G, Harris JF. 1975. Compositional Changes in Developing Rape Seed (Brassica-Napus L) *Planta* 123, 163-174.

O'Grady J, Schwender J, Shachar-Hill Y, Morgan JA. 2012. Metabolic cartography: experimental quantification of metabolic fluxes from isotopic labelling studies. *Journal of Experimental Botany* 63, 2293-2308.

Obendorf RL, Rytko GT, Byrne MC. 1982. Soya Bean Seed Growth and Maturation by In Vitro Pod Culture. *Ann Bot* 51, 217-227.

Obendorf RL, Rytko GT, Byrne MC. 1983. Soya Bean Seed Growth and Maturation by In Vitro Pod Culture. *Ann Bot* 51, 217-227.

Pandurangan S, Pajak A, Molnar SJ, et al. 2012. Relationship between asparagine metabolism and protein concentration in soybean seed. *Journal of Experimental Botany* 63, 3173-3184.

Paula Alonso A, Dale VL, Shachar-Hill Y. 2010. Understanding fatty acid synthesis in developing maize embryos using metabolic flux analysis. *Metabolic Engineering* 12, 488-497.

Periappuram C, Steinhauer L, Barton DL, Taylor DC, Chatson B, Zou J. 2000. The plastidic phosphoglucomutase from Arabidopsis. A reversible enzyme reaction with an important role in metabolic control. *Plant Physiology* 123, 1197-1197.

Pipolo AE, Sinclair TR, Camara GMS. 2004. Protein and oil concentration of soybean seed cultured in vitro using nutrient solutions of differing glutamine concentration. *Annals of Applied Biology* 144, 223-227.

Ratcliffe RG, Shachar-Hill Y. 2001. Probing plant metabolism with NMR. *Annual Review of Plant Physiology and Plant Molecular Biology* 52, 499-526.

Ratcliffe RG, Shachar-Hill Y. 2006. Measuring multiple fluxes through plant metabolic networks. *Plant Journal* 45, 490-511.

Rolletschek H, Radchuk R, Klukas C, Schreiber F, Wobus U, Borisjuk L. 2005. Evidence of a key role for photosynthetic oxygen release in oil storage in developing soybean seeds. *New Phytologist* 167, 777-786.

Rontein D, Dieuaide-Noubhani M, Dufourc EJ, Raymond P, Rolin D. 2002. The metabolic architecture of plant cells - Stability of central metabolism and flexibility of anabolic pathways during the growth cycle of tomato cells. *Journal of Biological Chemistry* 277, 43948-43960.

Ruuska SA, Schwender J, Ohlrogge JB. 2004. The capacity of green oilseeds to utilize photosynthesis to drive biosynthetic processes. *Plant Physiology* 136, 2700-2709.

Salon C, Munier-Jolain NG, Duc G, Voisin AS, Grandgirard D, Larmure A, Emery RJN, Ney B. 2001. Grain legume seed filling in relation to nitrogen acquisition: A review and prospects with particular reference to pea. *Agronomie* 21, 539-552.

Saravitz CH, Raper CD. 1995. Responses to Sucrose and Glutamine by Soybean Embryos Growth *In-Vitro*. *Physiologia Plantarum* 93, 799-805.

Schmidt MA, Tucker DM, Cahoon EB, Parrott WA. 2005. Towards normalization of soybean somatic embryo maturation. *Plant Cell Reports* 24, 383-391.

Schmutz J, Cannon SB, Schlueter J, et al. 2010. Genome sequence of the palaeopolyploid soybean. *Nature* 463, 178-183.

Schnarrenberger C, Oeser A, Tolbert NE. 1973. Two isoenzymes each of glucose-6-phosphate dehydrogenase and 6-phosphogluconate dehydrogenase in spinach leaves. *Archives of Biochemistry and Biophysics* 154, 438-448.

Schwender J. 2008. Metabolic flux analysis as a tool in metabolic engineering of plants. *Current Opinion in Biotechnology* 19, 131-137.

Schwender J, Goffman F, Ohlrogge JB, Shachar-Hill Y. 2004. Rubisco without the Calvin cycle improves the carbon efficiency of developing green seeds. *Nature* 432, 779-782.

Schwender J, Ohlrogge JB. 2002. Probing in vivo metabolism by stable isotope labeling of storage lipids and proteins in developing Brassica napus embryos. *Plant Physiology* 130, 347-361.

Schwender J, Ohlrogge JB, Shachar-Hill Y. 2003. A flux model of glycolysis and the oxidative pentosephosphate pathway in developing Brassica napus embryos. *Journal of Biological Chemistry* 278, 29442-29453.

Schwender J, Shachar-Hill Y, Ohlrogge JB. 2006. Mitochondrial metabolism in developing embryos of Brassica napus. *Journal of Biological Chemistry* 281, 34040-34047.

Shachar-Hill Y. 2002. Nuclear magnetic resonance and plant metabolic engineering. *Metabolic Engineering* 4, 90-97.

Sinclair TR, de Wit CT. 1975. Photosynthate and Nitrogen Requirements for Seed Production by Various Crops. *Science* 189, 565-567.

Smeekens S. 2000. Sugar-induced signal transduction in plants. *Annual Review of Plant Physiology and Plant Molecular Biology* 51, 49-81.

Smith AJ, Rinne RW, Seif RD. 1989. Phosphoenolpyruvate Carboxylase and Pyruvate Kinase Involvement in Protein and Oil Biosynthesis during Soybean Seed Development. *Crop Sci.* 29, 349-353.

Sriram G, Fulton DB, Iyer VV, Peterson JM, Zhou R, Westgate ME, Spalding MH, Shanks JV. 2004. Quantification of compartmented metabolic fluxes in developing soybean embryos by employing biosynthetically directed fractional ^{13}C labeling, two-dimensional [^{13}C, ^{1}H] nuclear magnetic resonance, and comprehensive isotopomer balancing. *Plant Physiology* 136, 3043-3057.

Sriram G, Fulton DB, Iyer VV, Peterson JM, Zhou RL, Westgate ME, Spalding MH, Shanks JV. 2004. Quantification of compartmented metabolic fluxes in developing soybean embryos by employing Biosynthetic ally directed fractional C-13 labeling, C-13, H-1 two-dimensional nuclear magnetic resonance, and comprehensive isotopomer balancing. *Plant Physiology* 136, 3043-3057.

Sriram G, Fulton DB, Shanks JV. 2007. Flux quantification in central carbon metabolism of Catharanthus roseus hairy roots by C-13 labeling and comprehensive bondomer balancing. *Phytochemistry* 68, 2243-2257.

Sriram G, Iyer VV, Fulton DB, Shanks JV. 2007. Identification of hexose hydrolysis products in metabolic flux analytes: A case study of levulinic acid in plant protein hydrolysate. *Metabolic Engineering* 9, 442-451.

Stepansky A, Leustek T. 2006. Histidine biosynthesis in plants. *Amino Acids* 30, 127-142.

Stephanopoulos G. 1999. Metabolic Fluxes and Metabolic Engineering. *Metabolic Engineering* 1, 1-11.

Stephanopoulos G. 2002. Metabolic engineering: Perspective of a chemical engineer. *Aiche Journal* 48, 920-926.

Stephanopoulos G, Vallino JJ. 1991. Network Rigidity and Metabolic Engineering in Metabolite Overproduction. *Science* 252, 1675-1681.

Stephanopoulos GN, Aristidou AA, Nielsen J. 1998. Metabolic Flux Analysis. *Metabolic Engineering*. San Diego: Academic Press, 309-351.

Stombaugh SK, Jung HG, Orf JH, Somers DA. 2000. Genotypic and environmental variation in soybean seed cell wall polysaccharides. *Crop Science* 40, 408-412.

Stombaugh SK, Orf JH, Jung HG, Somers DA. 2003. Relationships between soybean seed cell wall polysaccharides, yield, and seed traits. *Crop Science* 43, 571-576.

Sweetlove LJ, Fell D, Fernie AR. 2008. Getting to grips with the plant metabolic network. *Biochem J* 409, 27-41.

Szyperski T. 1995. Biosynthetically directed fractional ^{13}C-labeling of proteinogenic amino acids. An efficient analytical tool to investigate intermediary metabolism. *European Journal of Biochemistry* 232, 433-448.

Szyperski T. 1995. Biosynthetically Directed Fractional C-13-Labeling of Proteinogenic Amino-Acids - An Efficient Analytical Tool to Investigate Intermediary Metabolism. *European Journal of Biochemistry* 232, 433-448.

Thompson JF, Madison JT, Muenster AME. 1977. *In Vitro* Culture of Immature Cotyledons of Soybean (*Glycine-Max-L-Merr*) *Annals of Botany* 41, 29-39.

Troufflard S, Roscher A, Thomasset B, Barbotin JN, Rawsthorne S, Portais JC. 2007. In vivo C-13 NMR determines metabolic fluxes and steady state in, linseed embryos. *Phytochemistry* 68, 2341-2350.

van der Werf MJ, Overkamp KM, Muilwijk B, Coulier L, Hankemeier T. 2007. Microbial metabolomics: Toward a platform with full metabolome coverage. *Analytical Biochemistry* 370, 17-25.

van Winden W, Schipper D, Verheijen P, Heijnen J. 2001. Innovations in generation and analysis of 2D C-13, H-1 COSYNMR spectra for metabolic flux analysis purposes. *Metabolic Engineering* 3, 322-343.

van Winden WA, Heijnen JJ, Verheijen PJT, Grievink J. 2001. A priori analysis of metabolic flux identifiability from C-13-labeling data. *Biotechnology and Bioengineering* 74, 505-516.

Varma A, Palsson BO. 1994. Stoichiometric Flux Balance Models Quantitatively Predict Growth and Metabolic By-Product Secretion in Wild-Type *Escherichia-Coli W3110*. *Applied and Environmental Microbiology* 60, 3724-3731.

Wakao S, Andre C, Benning C. 2008. Functional analyses of cytosolic glucose-6-phosphate dehydrogenases and their contribution to seed oil accumulation in Arabidopsis. *Plant Physiology* 146, 277-288.

Walker DR, Parrott WA. 2001. Effect of polyethylene glycol and sugar alcohols on soybean somatic embryo germination and conversion. *Plant Cell Tissue and Organ Culture* 64, 55-62.

Weigelt K, Kuster H, Radchuk R, Muller M, Weichert H, Fait A, Fernie AR, Saalbach I, Weber H. 2008. Increasing amino acid supply in pea embryos reveals specific interactions of N and C metabolism, and highlights the importance of mitochondrial metabolism. *Plant Journal* 55, 909-926.

Wiechert W. 2001. C-13 metabolic flux analysis. *Metabolic Engineering* 3, 195-206.

Wiechert W, Mollney M, Petersen S, de Graaf AA. 2001. A universal framework for C-13 metabolic flux analysis. *Metabolic Engineering* 3, 265-283.

Wilcox JR, Shibles RM. 2001. Interrelationships among seed quality attributes in soybean. *Crop Science* 41, 11-14.

Williams TCR, Miguet L, Masakapalli SK, Kruger NJ, Sweetlove LJ, Ratcliffe RG. 2008. Metabolic network fluxes in heterotrophic arabidopsis cells: stability of the flux distribution under different oxygenation conditions. *Plant Physiology* 148, 704-718.

Willker W, Flogel U, Leibfritz D. 1997. Ultra-high-resolved HSQC spectra of multiple-C-13-labeled biofluids. *Journal of Magnetic Resonance* 125, 216-219.

Wilson LA. 1995. Soy foods. In D. R. Erickson (ed.) Practial handbook of soybean processing and utilization. AOCS Press, Champaign, IL and United Soybean Board, St. Louis, MO., 428-459.

Wuthrich K. 1976. NMR in Biological Research: Peptides and Proteins. *North Holland, Amsterdam*.

CHAPTER 4

SOYBENA SOMATIC EMBRYOS METABOLIC FLUX MAP COMPARISONS OF PLASTIC PHOSPHOGUCOMUTASE KNOCK OUT AND DIFFERENT CARBON NUTRITIONAL SUPPLEMENT IN THE MEDIUM

Quyen Truong[1], John D. Everard[2], Jacqueline V. Shanks[1]

[1]Department of Chemical and Biological Engineering, Iowa State University, Ames, IA 50010, USA;[2]DuPont Agricultural Biotechnology Research and Development, DuPont Experimental Station, Wilmington, Delaware 19880, USA

Abstract

Soybean (*Glycine max*) seed provide significant amounts of protein and oil level for the food and feed industries. The objective of this project is to use metabolic flux mapping to gain an understanding of resource partitioning and metabolic control points of protein and oil biosynthesis in soybean (*Glycine max)* seeds, using Soybean Histodifferentiation and Maturation (SHaM) embryos as the model system. SHaM embryos of transgenic cultures with the plastidic phosphoglucomutase (PGM) gene knocked out (PGM-KO), and the control (PGM-null) are cultured in sucrose concentrations ranged from 88 to 234 mM as a carbon source and initial glutamine concentrations ranged from 20 to 60 mM as a nitrogen source. These concentrations correspond to C:N ratios ranging from 8.8 to 70.2. Two C: N mole ratio conditions are further examined through metabolic flux analysis with labeling experiment of U-$^{13}C_{12}$ sucrose for both PGM culture. The result indicates that: (1) protein and oil of PGM-KO were consistently higher than the PGM-null; (2) content in PGM-KO shows nearly two fold as compared to PGM-null; and (3) for both PGM culture,

protein content is strongly correlated with the glutamine uptake rate. Fluxes through cytosolic glucose-6-phosphate isomerase, transketolase, and transaldolase, contributed significantly to the soluble sugar content for PGM-KO culture. These fluxes changed in response to the absence of starch synthesis. Additionally, flux toward malate transporter provided significantly to the oil and protein synthesis. These results and high reproducibility in the data indicate that soybean somatic embryos in SHaM media are a excellent model system for hypothesis testing to provide insight into the metabolic process regulating resource partitioning in soybean seed.

1 Introduction

Soybean [*Glycine max* (L.) Merr.] seeds are a source of food, feed, and fuel for economically valuable products in many industries. Soybean seed composition at maturity is 40% protein and 20% extractable oil by dry weight (Sinclair and de Wit, 1975). However, one of the dilemmas is that the oil and protein contents in soybean seeds appear to be inversely related (Hartwig and Kilen, 1991) with an approximate trade-off of 1% of oil for every 2% of protein (Clemente and Cahoon, 2009). Therefore, it is desirable to increase the oil content of soybeans at the expense of the 40% remaining of the biomass of seed (i.e., starch and soluble carbohydrate) and keep protein content at the level that meets market demands. To accomplish this goal, it is crucial to have a comprehensive understanding of the control points the partitioning of carbon and nitrogen between oil, protein, starch, and soluble carbohydrates pools. Our current knowledge in the seed development indicates that carbon and nitrogen, especially sucrose and amino acids, supplies into the storage compounds (Baud and

Lepiniec, 2010; Meyer and Kinney, 2009). Furthermore, flux analysis with cultured developing seeds has provided insight into the genes and metabolic pathways involved in the regulation of storage reserve synthesis in plant model systems (Allen *et al.*, 2009; Allen and Young, 2013; Alonso *et al.*, 2011; Hay and Schwender, 2011; Iyer *et al.*, 2008; Junker *et al.*, 2007; Schwender *et al.*, 2004; Schwender and Ohlrogge, 2002; Schwender *et al.*, 2003; Sriram *et al.*, 2004). Metabolic flux analysis on microorganisms is routinely exercised to examine the effect of genetic manipulation on their central metabolisms (Sauer, 2006). To our knowledge, the first report is in the plant model Arabidopsis, documenting the use of metabolic flux analysis to investigate well-characterized genetic mutants and to make rational transgenic manipulation to improve seed quality (Lonien and Schwender, 2009). However, employing metabolic flux analysis on a transgenic soybean crop can shed light on the regulation of central carbon metabolism leading to protein and oil. Such information could enhance the chances of successfully engineering soybean seeds to improve their quality and economic value.

In this study, we use soybean somatic transgenic embryos cultured in Soybean Histodifferentiation and Maturation (SHaM) media. SHaM embryos are perturbed in starch synthesis as a method to gain an understanding of resource partitioning and metabolic control points of protein and oil biosynthesis in soybean seeds by metabolic flux analysis. With this objective, we adapt a method of using SHaM embryos cultures previously established as an alternative to soybean zygotic embryos (Iyer *et al.*, 2008; Sriram *et al.*, 2004). SHaM embryos have many advantages as a model system to understand the metabolism of developing zygotic embryos (seed). They are

compositionally and metabolically similar to zygotic embryos, and can provide stable transgenic cultures for experimentation directed at understanding the genotype-phenotype relationship (Nishizawa *et al.*, 2010). Furthermore, SHaM embryos have proven to be an excellent system for the discovery of genes leading to oil content increases in mature soybean seed (Meyer *et al.*, 2012). Perhaps the greatest advantage provided by SHaM embryos is that they can be generated and examined more quickly than zygotic embryos. The cycle time from transformation-to-harvest of SHaM embryos is approximately eight to ten weeks, at which time stable transgenic events can be identified and maintained for experimentation. In contrast, it takes approximately 11 months to generate seed-bearing transgenic plants, and then several more generations are required to obtain homozygous plants capable of generating zygotic embryos for study. It is also possible to initiate SHaM embryos from stably transformed soybean plants. Therefore, SHaM embryos of transgenic cultures with the plastidic phosphoglucomutase (PGM) gene knocked out (PGM-KO), and the control (PGM-null) are chosen for the flux studies. The PGM-KO transgenic soybean lines, in which PGM has been severely down regulated, but not entirely eliminated, was incapable of producing substantial starch during development (Allen et al., 2008). SHaM embryos generated from these steady transformed lines (PGM-KO) and from similarly transformed lines that do not express this trait (PGM-null) can be maintained in culture and used for experimentation.

Steady state ^{13}C-metabolic flux analysis (MFA) has emerged as a powerful set of techniques for developing a deep understanding of metabolic control at the cellular level (Kruger *et al.*, 2012; Libourel and Shachar-Hill, 2008; O'Grady *et al.*, 2012;

Ratcliffe and Shachar-Hill, 2006; Schwender, 2008; Stephanopoulos and Vallino, 1991; Wiechert, 2001). Steady state ^{13}C-MFA utilizes patterns of labeled metabolites, resulting from isotopic labeling of carbon experiments, as well as input and output measurements to establish internal metabolite flux values for generating detailed flux maps. Nodal analysis of the flux map models can be used to identify differences between genetic and environmental variants (Stephanopoulos and Vallino, 1991), and can provide an experimental framework for investigating differences between transgenic culture and its control under different environmental perturbations at the metabolic level. A novel generic ^{13}C metabolic flux analysis (^{13}C MFA) tool, NMR2Flux model, was developed for zygotic soybean embryos (Sriram *et al.*, 2004) and was recently used to identify the impact of temperature on carbon partitioning into protein and oil in zygotic embryos (Iyer *et al.*, 2008). In this study, we apply ^{13}C MFA to transgenic SHaM embryos to develop a better understanding of carbon partitioning in oil and protein. The soybean transgenic (PGM-KO) embryos cultured in two levels of sucrose concentrations are compared with their respective control (PGM-null) embryos. The result of this study provides insight in (1) how eliminating starch affect the other storage composition, (2) how the levels of carbon in the media affect the transgenic and its control embryos. Hypotheses suggested by metabolic flux analysis studies can then be further tested by genetic manipulation of the SHaM embryos in a semi-empirical metabolic engineering design strategy.

2 Materials and Methods

2.1 Plant material and culture conditions

Cultures of transgenic proliferative soybean somatic embryos with the plastidic phosphoglucomutase (PGM) gene knocked out (PGM-KO), and the control (PGM-null) were provided by DuPont Pioneer Hi-Bred, and were maintained by weekly subculture, in SB196 medium. SB196 medium is a modified form of MSD20 medium (Walker and Parrott, 2001), and contained FN Lite Halides, FN Lite P, B, Mo, Murashige-Skoog (MS) sulfate, MS Fe EDTA, B5 vitamins, KNO_3, L-asparagine, and 29 mM sucrose. In addition, the SB196 medium contained a moderately high concentration of synthetic auxin, 2,4-dichlorophenoxyacetic acid (10 mg/l). Clusters of globular-stage embryos (approximately 10 mg) were used to initiate cultures by dropping them into 250-ml Erlenmeyer flasks containing 35 ml SB228 liquid media (Schmidt *et al.*, 2005). The SB228 medium contained FN Lite macro salts, MS micro salts, B5 vitamins, $CaCl_2$, L-methionine, 88 mM sucrose, 30 mM glutamine, and 165 mM sorbitol. Embryos were induced, maintained, and matured under a light intensity of 35 – 50 µmol photons m^{-2} s^{-1}, provided by cool-white fluorescent bulbs, for a 16-h photoperiod. PGM cultures were maintained at 26 °C and were continuously shaken at 130 rpm for two weeks to build biomass prior to experimentation.

After two weeks in culture, working in a laminar flow hood, clusters of PGM-KO and PGM-null embryos collected from 20 flasks derived from the same subculture cycle, were separated into batches of approximately 40 to 50 randomly chosen, uniformly sized embryos, with an initial fresh weight of 1.00 g (measured and

recorded to a precision of 0.01g). These were placed into 250-ml Erlenmeyer flasks containing 35 ml of SHaM media with different sucrose and glutamine concentrations to be tested (Table 1). For labeling experiments, the two SHaM media conditions were: (1) with 146 (10% Uniform-13C labeled sucrose from Isotec, Miamisburg, OH, USA) and 37 mM glutamine (which represents the carbon and nitrogen sources for the C: N mole ratio of 24), (2) 234 mM sucrose (10% Uniform-13C labeled sucrose) and 37 mM glutamine (which indicates the carbon and nitrogen sources for the C: N mole ratio of 38). The media for each flask was exchanged every three days. Cultures were harvested after six days by pouring the contents of each culture into a sieve. The harvested embryos were rinsed with 100 ml distilled water, blotted dry, weighed (fresh weight determination) and frozen in liquid nitrogen prior to lyophilization at -50°C and 0.0158 mbar for 72 hours. The lyophilized embryos were weighed (dry weight determination) and were finely ground using a Geno/Grinder® (SPEX SamplePrep, Metuchen NJ, USA) prior to further analysis.

2.2 Relatively Growth Rate

The growth rates of the PGM-KO and PGM-null embryos were normalized based on their relative linear growth rates. Growth was expressed as the dry mass added during a given growth interval, per unit of original dry mass at the beginning of the growth interval. The relative growth rate (day^{-1}) was calculated using the following equation:

$$\mu_{relative} = \frac{\left[\frac{X_{Dayn} - X_{Day0}}{t}\right]}{X_{Day0}}$$

where, X_{Dayn} is the tissue dry-weight at final harvest , X_{Day0} is the initial dry-weight of the embryo, described below, and t is the culture duration. To determine the initial dry weight of the culture, three 1.00 gram fresh weight batches of representative embryos were prepared at the same time that the experimental cultures were being initiated (see above). The embryos were placed into pre-weighed 50-ml centrifuge tubes and were frozen in liquid nitrogen prior to lyophilization, as described above. After lyophilization, the embryo dry weight was determined and used in the calculations described above.

Absolute flux values were normalized to relative fluxes based on sucrose uptake rate of 100 moles being accumulated in the biomass. This normalization is equal to the total biomass accumulation rate, and therefore makes it easy to compare the flux across the phenotypes. For example, the absolute flux through the plastidic glycolytic pathway, hxi^p, was 2000 µmol g-Dwt^{-1} day^{-1}, and the corresponding sucrose intake rate for that culture was 1000 µmol g-Dwt^{-1} day^{-1}. Therefore, the relative flux through this pathway after normalization for the sucrose intake of 100 moles was [(2000/1000)*100 = 200], the flux for hxi^p would be 200.

2.3 Biomass Quantification and Fractionation

The dry weight percents of carbon, hydrogen, and nitrogen were measured by elemental analysis using the Perkin-Elmer Model 2400 Series II CHN&S Elemental analyzer (Chemical Instrument Facility, Chemistry Department at Iowa State University). In the determination of biomasses used throughout this study, total protein content was calculated from the elemental analysis (% N x 6.25).

Approximately 100 mg of dry powdered embryo sample was extracted with 1 ml of n-hexane at 40°C for 1 hour, then the extracts were centrifuged at high speed (13,200 rpm) for 10 minutes at room temperature. The process was repeated five times and after each extraction, the lipid-containing solvent was collected in separate pre-weighed glass tubes, and dried for four days in a hood at room temperature. The mass of lipids after solvent evaporation were measured gravimetrically. The remaining hexane-extracted biomass was dried and further extracted for protein in 600 µl of 200 mM phosphate buffer (pH = 7.2) containing 14 mM β-mercaptoethanol at 4°C for 20 minutes to dissolve and suspend the protein into solution. The extracts were centrifuged at high speed at 4°C for 15 minutes, and the supernatant was transferred into a micro-centrifuge tube. The extraction was performed on the pellet two more times with 400 µl of the buffer (per extraction) and the supernatants were pooled with that from the initial extraction (approximately 1.2 ml final volume). Protein contents of the labeling experiment extracts were measured by Bradford Assay (Bradford, 1976) to obtain total extracted protein content prior to dialysis and acid hydrolysis of the protein sample.

Residual soluble sugars in the defatted/deproteinated- biomass pellets were extracted into 1 ml of 80% aqueous ethanol in a water bath sonicator (Fisher Scientific FS110H Ultrasonic Cleaner) at 60°C for 20 minutes. The extraction was repeated four times, and after each extraction, the solvent containing the soluble sugars were collected in separate pre-weighed glass tubes. The remaining pellets from each sample were placed in 1.5 ml of DDI water, covered with foil and autoclaved (liquid cycle at 121 °C and a pressure of 15 psi) for 30 minutes prior to

starch digestion and extraction. Starch was digested in 1.5 ml of 100 mM citrate buffer (pH 5.0) containing amyloglucosidase at a ratio of 0.025 mg enzyme: 1 mg of tissue dry weight (approximately 2 µl of amyloglucosidase or 0.6 units of enzyme per sample). Samples were incubated overnight in a 30°C water bath. Starch content was quantified using a Glucose Assay Kit (Sigma, St. Louis, MO).

Ash content was determined by complete combustion of the samples in a Thermo-gravimetric Analyzer (TGA 7) running the following temperature ramp program: ramp from room temperature to 100 °C and hold for 10 min; ramp to 200 °C at 10 °C/min; hold at 200 °C for 10 min; ramp to 400°C at 10°C/min; hold at 400 °C for 30 min; ramp to 700 °C at 10 °C/min; hold at 700 °C for 180 min; drop to, and hold at, room temperature until sample removal. The remaining residue in the crucible was taken as the ash content. DNA/RNA content was assumed to be 5 percent on a dry weight basis (Stephanopoulos *et al.*, 1998).

The percentage of the residual biomass fraction was estimated by subtraction of the measured sum of the biomass, i.e., protein (by combustion analysis) + oil + starch + soluble sugar + ash and an estimation of the DNA/RNA content (see above), from the original mass of tissue extracted. Ash content was determined by Thermo-Grametric Analysis (TGA) as described previously. DNA/RNA content was assumed to be 5 percent on a dry weight basis (Stephanopoulos *et al.*, 1998).

Sucrose and glutamine consumption rates from the media were determined by measuring the sucrose and glutamine concentrations in the culture media at the initiation of the experiments, at the three-day exchange, and at final harvest. Each media sample was filtered with a syringe filter (25mm, 0.45 µm, Nylon,

XPERTEX®) prior to injection into an HPLC (Waters 1525 Binary Pump and 717^{plus} AutoSampler fitted with an Aminex HPX-87K column, and a Waters 2414 RI detector). Samples were eluted with a water mobile phase at 50°C, at a flow rate of 0.35 ml/minute; each run was 80 minutes. By running the separation for 80 minutes, both sucrose and glutamine were detected in the same HPLC run. The HPLC run also detected sorbitol in the liquid medium. A series of sucrose (15, 44, 88, 146, 176, and 234 mM), glutamine (5, 10, 20, 30, 40, and 60 mM), and sorbitol (50, 100, 150, 200 mM) standards were prepared and measured with the unknown samples. Standard curves were generated by plotting the area under the peaks versus their concentrations, and were used to compute the concentrations of the unknown samples.

Sucrose and glutamine consumption rates were calculated using the following equation:

$$\frac{dS}{dt} = \frac{\left(S_{Day6} - S_{Day0}\right) \times V}{\left(t_{Day6} - t_{Day0}\right) \times \left(X_{Day6} - X_{Day0}\right)}$$

where S is the media sucrose or glutamine concentration at the onset and end of the culture period, V is the volume of the media after three or six-days of culture, t is time, and X is the embryo dry weight.

2.4 Protein/Starch Hydrolysis and NMR Sample Preparation for labeling PGM cultures

Protein extracts from labeling experiments of PGM-KO 24, PGM-KO 38, PGM-null 24, and PGM-null 38, in total of 12 cultures, were dialyzed using dialysis cassettes (Pierce, Rockford, IL) in 10 mM phosphate buffer (pH 7.2) at 4 °C for 24 hours to remove any soluble sugar residues in the sample. The protein and starch

samples were vacuum hydrolyzed in hydrolysis tubes (Pierce Endogen, Rockford, IL) using 6Nconstant boiling hydrochloric acid (HCl) (Pierce, Rockford, IL). The acid was added in the ratio of 6 ml HCl:10 mg protein/starch. However, there were no starch samples for the PGM-KO cultures. The sample tubes were sealed and gently vortexed for 1 minute. The hydrolysis tubes containing the protein and starch samples were evacuated first with the vacuum, and then purged with nitrogen gas. The process was repeated three times, and finally the samples were re-evacuated to prevent oxidation during hydrolysis. The 12 protein and 6 starch samples were hydrolyzed for 4 hours at 140-145°C in a heat block. After hydrolysis, a Rapidvap evaporator (Labconco, Kansas City, MO) was used to evaporate the hydrochloric acid for 2 hours at 35 % speed, 40 bar pressure, and a temperature of 45°C. Each dried hydrolysate (one for protein and one for starch) was reconstituted in 2 ml of de-ionized water using a Rapidvap evaporator for 30 minutes at 70 % speed, atmospheric pressure, and room temperature. The reconstituted samples were filtered using a centrifuge and Spin-X® centrifuge tube with filter. The filters were discarded, the protein and starch sample tubes were tightly sealed with parafilm, and holes were poked through the parafilm to allow water vapor to escape during lyophilization. The protein and starch samples were frozen in a -80°C freezer for 24 hours and lyophilized at -50°C and 0.0158 mbar for 72 hours or until the samples were completely dry. The protein and starch samples were reconstituted in 600 µl of deuterium oxide and vortexed. Then, a small portion of the protein sample was diluted to 20-50 µl using DDI water and analyzed by HPLC for proteinogenic amino acid content. Sixteen proteinogenic amino acids can be analyzed by using pre-

column derivatization with 6-Aminoquinolyl-N-Hydroxysuccinimidyl carbamate (Cohen, 2000). The remainders of the protein and starch samples were reconstituted with deuterium oxide as described above; the sample pH was adjusted to 0.8-1.0 using deuterium chloride. A pinch (2 or 3 flakes) of 3-(Trimethylsilyl)-1-propanesulfonic acid sodium salt (TMS) internal standard flakes was added. The samples were then thoroughly mixed with a vortex mixer, and were transferred to clean NMR tubes prior to analysis. NMR samples were submitted to NMR facility (Department of Biochemistry, Biophysics, and Molecular Biology at Iowa State University).

2.5 NMR spectroscopy and Analysis for PGM cultures

2-Dimensional [^{1}H, ^{13}C] Heteronuclear Single Quantum Correlation (HSQC) spectroscopy of hydrolyzed 12 protein and 6 starch samples was performed on a Bruker Avance DRX 500 MHz spectrometer at 298 K. The HSQC spectra were acquired using a modified version of the INEPT spectra (insensitive nuclei enhanced by polarization transfer) pulse sequence (Bodenhausen and Ruben, 1980). TMS is used as an internal standard: the reference zero p.p.m. was set by the methyl signal of TMS. The magnetic resonance frequency of ^{13}C (F1 dimension) and ^{1}H (F2 dimension) are 125.7 MHz and 499 MHz, respectively. The spectral widths along the F1 dimension and F2 dimensions are 5028.05 Hz and 5482.26 Hz, respectively. Peak aliasing was applied to minimize the sweep width in the F1 dimension. The number of complex data points collected were 900 (^{13}C) and 1024 (^{1}H). The number of scans ranged from 16 to 32.

NMR spectra were acquired using Xwinnmr (Bruker) software and analyzed with NMRview freeware (Johnson and Blevins, 1994) using the Linux operating system. NMR Cross-peaks signals of singlet “s”, doublet “d”, and triplet “t” can be processed directly to obtain raw signal intensities using the NMRview software. Both aliphatic and aromatic carbon atoms of 16 proteinogenic amino acids in the spectra were identified in the hydrolysates of the soybean protein samples extracted from the zygotic and somatic embryos (Sriram *et al.*, 2004). Each carbon of the proteinogenic amino acids was detected by its unique ^{13}C and ^{1}H chemical shifts (Wuthrich, 1976) as well as the distinctive coupling patterns and J-coupling constants (Jcc) (Harris, 1983). The chemical shifts and J-coupling-constant (Jcc) values were obtained from Krivdin and Kalabin, 1989. Further, to quantify overlapping multiplets [(s, d1, s2, dd), (s, d1, d2, d3, dd1, dd2, dd3, qd)] on the 2D [^{1}H, ^{13}C] HSQC spectra, peak de-convolution software, developed by Sriram et al., (2004), based on a spectral model originally proposed by van Winder et al., (2001), was used. Since the peaks of the aromatic amino acids and some of the protein-associated glucosyl units were crowded, additional 2-D HSQC spectra were acquired that when J-scaled along the ^{13}C dimension could be resolved, by integrating pulse sequences reported previously (Brown, 1984; Willker *et al.*, 1997). J-scaling increases multiplet separation by an even integral factor J and eliminates multiplet overlap. J-scaling factors of six were employed in processing aromatic amino acid and some of glucosyl units (Sriram *et al.*, 2007).

2.6 Metabolic Flux Analysis: Reaction Network for Central Carbon Metabolism

The reaction network of central carbon metabolism of soybean embryos was modeled, consisting of reactions of glycolysis (GLY), oxidative pentose phosphate pathway (OxPPP), tricarboxylic acid cycle, and anaplerotic reactions. A total of 76 metabolic reactions were formed by carbon arrangements, and the model simulated isotopomers with carbon labeling in 135 metabolic pools (Supplement Table S1). The metabolic flux map in this study (Fig. 7) is based on a flux model of developing zygotic embryos (Sriram *et al.*, 2004) with extensive modification from additional measurements and transcriptomic data for both zygotic and somatic embryos (data not shown). For clarification, in the cytosol and plastid oxidative pentose phosphate, we combined the five-carbon atom metabolites ribose 5-phosphate, ribulose-5-phosphate and xylulose 5-phosphate into one metabolite called pentose-5-phosphate (P5P). In the cytosol and plastid glycolysis, we also combined the three-carbon atom metabolites dihydroxyacetone phosphate, glyceraldehydes-3-phosphate, phosphoenolpyruvate and pyruvate into one metabolite called triose-3-phosphate (T3P). Due to similar carbon atom rearrangement and rapid equilibration of these metabolites (van Winden *et al.*, 2001), combining metabolites is supported. The NMR data derived from the glucosyl units associated with the protein and starch pools provide crucial information about fluxes through the cytosolic and plastidic glucose-6-phosphate pools, respectively. This information can be used to estimate carbon fluxes through the GLY and OxPPP pathways in the two separate cellular compartments (Sriram *et al.*, 2007). For more detail of the reaction network representing soybean central metabolism, the reader should refer to previous work

(Sriram *et al.*, 2004). Approximately 500 simulations (each time with random start values from the free fluxes) were performed for 12 PGM cultures in order to ensure that a global optimum had been located. The error in the isotopomer measurements was used to perform random Monte Carlo simulations to generate probability distributions for the flux. Hence, 12 flux maps were generated from the both PGM cultures with the two different C: N mole ratios.

2.7 Statistical Analysis

All analyses were performed in triplicate. Biomass composition, growth rate, and substrates consumption rate data shown in figures represent the means of three determinations ± standard errors. The means were statistically compared by comparison for all pairs using Tukey's test. Means within each culture with the same letter are not significantly different. The statistical calculations were performed using the JMP software (v. 9.0.1, SAS Institute Inc. Cary NC, USA) and all statistical significance test used $\alpha = 0.05$.

Each PGM culture experiment was performed with three biological replicates (n = 3) and for two conditions of C: N ratios for each PGM culture, for a total of 12 cultures. As mentioned above, 500 simulations were computed for each PGM culture to develop a single metabolic flux map for each culture. Average was determined between three biological replicates, but the standard error was calculated by pooled variance method with the equation below:

$$S_p^2 = \frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2 + (n_3 - 1)S_3^2}{n_1 + n_2 + n_3 - k}$$

where n = 500, S = standard error for each biological replicate, and k = the number of samples are being combined (in this case is 3).

Each flux reported in the table is an average of the biological triplicate ± Pooled variance. The means were statistically compared by comparison for all pairs using Tukey's test between the four cultures: PGM-KO 24, PGM-KO 38, PGM-null 24, and PGM-null 38. Means within each culture with the same letter are not significantly different as displayed in the figures. The statistical calculations were performed using the JMP software (v. 9.0.1, SAS Institute Inc. Cary NC, USA) and all statistical significance test used $\alpha = 0.1$.

3 Results

3.1 Effect of C: N mole ratio on Relative Growth Rates of PGM-KO and PGM-null

Application of metabolic engineering to soybean somatic embryos in SHaM media can be used to enhance the understanding of the protein-oil relationship. This study was designed to elucidate the influence of PGM on the partitioning of protein and oil under conditions of different carbon (sucrose) to nitrogen (glutamine) [C: N] mole ratios. The linear relative growth rates of transgenic soybean somatic embryos, PGM-null and PGM-KO under 20 different C: N ratios are shown in Figure 1. The supplemented sucrose concentration range (88 to 234 mM) and glutamine concentration range (20 to 60 mM) were selected based on published literature (Gifford and Thorne, 1985; Hsu *et al.*, 1984; Pipolo *et al.*, 2004; Saravitz and Raper, 1995; Schmidt *et al.*, 2005). Table 1 provides the ratio of C moles of sucrose per N moles of glutamine fed corresponding to initial concentrations of sucrose and glutamine in SHaM media. The relative growth rates of both PGM cultures were

calculated as mass added during a given growth interval per unit of original mass at the beginning of the growth interval, unit (day^{-1}). Figure 1 indicates the average relative growth rate of all media conditions of PGM-KO and PGM-null, 0.60 ± 0.08 (1/Day) and 0.62 ± 0.08 (1/Day) respectively, during 6-day culture; each media condition represents three flasks for biological replicates in triplicate. The result from PGM-KO embryos relative growth rate indicated that growth-inhibition may occur as the C: N mole ratio increases, when compared to the PGM-null embryos (Figure 1). The relative growth rate values of both PGM cultures were compared to *cv Jack* of soybean zygotic embryos of 0.52 ± 0.10 (day^{-1}) (data not shown). The error bars for relative growth rate were tighter in PGM-KO culture than in PGM-null culture, perhaps indicating that the PGM-KO culture was more stable. Moreover, the data validates that the relative growth rates of PGM-KO and PGM-null embryos in the range of sucrose and glutamine concentrations were constant and the embryos were not under carbon- and nitrogen-limited regimes.

3.2 Effect of C: N mole ratio on Biomass Composition of PGM-KO and PGM-null

The starch content in PGM-null and PGM-KO embryos were measured, as shown in Figures 2A and 2C, respectively. The data indicates that there was an insufficient amount of starch in the PGM-KO when compared to PGM-null embryos, as expected. The PGM transgenic event, created in Pioneer Hi-Bred soybean variety 93B92, severely down-regulated the plastidic PGM, essentially making the seed incapable of producing starch during development (Allen et al., 2008). In PGM-KO embryo, the starch content was 0.53 ± 0.10 % Dwt across the broad range of C:N mole ratios. These extremely low starch contents are consistent with results observed

previously for PGM-KO embryos cultured at the DuPont Experimental Station: for 2 week old PGM-KO embryos in SHaM media, 88 mM sucrose and 30 mM glutamine, after proliferation, the starch content was 0.25 ± 0.09 % Dwt. On the other hand, the PGM-null starch content increased with the increasing of C: N mole ratio, at a linear correlation coefficient $R^2 = 0.84$, Figure S1. This analogous result was discovered in soybean somatic embryos *cv Jack* in the same C:N mole ratio of the SHaM media (data not shown). To further investigate the effect of sucrose and glutamine concentrations on carbohydrates, soluble sugars of PGM cultures were extracted, and the results are shown in Figures 2B and 2D. No discernible correlation between the soluble sugars in the PGM embryos and C: N mole ratio is observed. However, PGM-KO embryos, when compared to PGM-null, consistently displayed higher soluble sugars content across the entire C:N mole ratio range, as illustrated in Figure B. On average, the soluble sugars content based on percentage dry weight of PGM-null and PGM-KO were 16.46 ± 0.64 and 28.92 ± 0.88 respectively. Therefore, the result shows that the soluble sugars content in PGM-KO was 80% higher, on average, than in PGM-null.

Protein content in dry weight (Dwt) percentage was analyzed for PGM-KO and PGM-null cultures (Figures 2A and 2C). The protein contents for both PGM cultures were analyzed based upon elemental analysis. The data from both PGM cultures indicate similar trends of increasing protein content with increasing glutamine and decreasing sucrose supply. This relationship is more clearly seen if the protein content is plotted against the C:N mole ratio, as shown in Figures 2A and 2C. The protein content of PGM-KO and PGM-null doubled as the C:N mole ratio decreased.

The protein content for both PGM cultures increased from ~16 to 40% Dwt as the C:N mole ratio decreased from 70.2 to 8.8. These results are consistent with the role of glutamine as a main reduced nitrogen source in enhancing the protein biosynthesis. Nevertheless, the general trend showed that the protein content of PGM-KO embryos was higher at each media condition tested, with an average Δ of 2.3 Dwt % and maximum Δ of 4 Dwt % when compared with PGM-null culture, where Δ = PGM-KO – PGM-null. This increase in protein content does not offset the decrease in starch content resulting from knocking out PGM (Figures 2B and 2D). To investigate whether the oil content varied at different concentrations of sucrose and glutamine in the media, hexane extraction was performed on both PGM cultures (Figures 2A and 2C). The range of oil content based on dry weight (Dwt) was 3 – 6% for PGM-KO and PGM-null embryos. The biosynthesis of oil has no correlation with the C: N mole ratios as was seen in the case of protein biosynthesis.

To complete the biomass composition balance, the residual biomass fraction was estimated based on mass balance for both PGM cultures. The percentage residual biomass fraction was performed by subtraction of the sum of the biomass composition (protein + oil + starch + soluble sugar + ash + DNA/RNA). Ash content was examined by Thermo-Gravimetric Analysis (TGA) as described in the Materials and Methods Section. DNA/RNA content was assumed to be 5 percentage of dry weight basis (Stephanopoulos et al., 1998). A similar trend was observed for PGM-null and PGM-KO in that the residual biomass fraction increased as C:N mole ratio increased, as shown in Figures 2C and 2D. The result illustrates that the C: N mole

ratio regime there is no significant difference statistically between the residual biomass fractions for PGM-KO and PGM-null.

The protein content of both PGM cultures indicated a strongly positive correlation with the glutamine consumption rate, Figures 3A and 3B, and data indicated linear correlation coefficients of PGM-null and PGM-KO embryos were $R^2 = 0.74$ and $R^2 = 0.82$, respectively. The glutamine consumption rate provides a clear indication that PGM-containing cultures consumed more glutamine than null cultures, as media glutamine concentrations increased from 20 to 60 mM. Measurements of endogenous free amino acids pools show clearly that the majority of the glutamine taken up by the embryos was converted to protein. However, the sucrose consumption rate was more complex in both PGM cultures. No significant correlation between sucrose consumption rate and the media C: N mole ratio was observed (data not shown). In our studies, the medium was replaced every three days to ensure that the nutrient supplies were constant for carbon and nitrogen sources throughout the experimental period. The percentages of sucrose depletion under the 20 media conditions were examined for day 3 and day 6 cultures and indicated that the percentages of sucrose depletion were less than 30% for both PGM cultures (data not shown).

3.3 ^{13}C labeling for PGM-null and PGM-KO cultures

Application of steady state ^{13}C-Metabolic flux analysis (MFA) of soybean somatic embryos in SHaM media on the plastidic phosphoglucomutase (PGM) gene knocked out (PGM-KO) and control (PGM-null) can be used to enhance the understanding of protein and oil relationship. This study was designed to elucidate the influence of PGM on the partitioning of protein and oil under two media

conditions with ^{13}C-Uniformly labeled sucrose. The two media conditions, with sucrose and glutamine concentrations as a major sources of carbon and nitrogen for PGM-KO and PGM-null embryos cultured *in vitro*, were designed to mimic conditions experienced by developing seed on the plant, and were set at (1) 146 mM sucrose (10% Uniform 13C-sucrose) and 37 mM glutamine, (2) 234 mM sucrose (10% Uniform 13C-sucrose) and 37 mM glutamine (Allen *et al.*, 2009; Allen and Young, 2013; He *et al.*, 2011; Hsu *et al.*, 1984; Iyer *et al.*, 2008; Pipolo *et al.*, 2004; Saravitz and Raper, 1995; Schmidt *et al.*, 2005). Table 1 shows the ratio of moles of C, provided by sucrose, to the moles of N, provided by glutamine in the initial SHaM media for both PGM cultures. Hence, the two media conditions correspond to C: N mole ratio of 24 and 38. Flux for the PGM knockout (PGM-KO) relative to the control (PGM-null), as well as carbon source in the media are shown to influence the major storage compounds in both cultures. Developing embryos of each genotype were cultured in the media conditions mentioned above for 6 days, with media exchanged every three days. After 6 days, the embryos were harvested, weighted, and lyophilized, the relative growth rates of both PGM cultures were calculated, and the major storage compounds of protein, oil, starch, soluble sugars were determined. Each media condition represents three flasks for biological replicates in triplicate, the relative growth rate based on dry weight of PGM-null and PGM-KO embryos cultured in C: N mole of ratio 24 for 6 d were 0.59 ± 0.01 day^{-1} and 0.62 ± 0.01 day^{-1}, respectively. Similar, for the media condition C: N mole ratio 38, the relative growth rates were 0.64 ± 0.01 day^{-1} and 0.60 ± 0.01 day^{-1}, PGM-null and PGM-KO

respectively. The error bars for relative growth rates of PGM cultures were similar and insignificant, perhaps indicating that both cultures were stable.

The starch content in PGM-KO and PGM-null embryos were measured, as shown in Figure 4B. The data indicated that there was an insufficient amount of starch in the PGM-KO when compared with PGM-null embryos, as expected. On the other hand, the PGM-null starch content was 12.34 ± 0.49 % Dwt with no significant difference between the two medias. To investigate the effect of sucrose concentrations on carbohydrates, soluble sugars of PGM cultures were extracted, and the results are shown in Figure 4F. PGM-KO embryos consistently displayed higher soluble sugars content for both media conditions when compared to PGM-null. Furthermore, at higher sucrose concentration in the media, both embryos displayed higher soluble sugars content. The soluble sugars content, based on percentage dry weight, compared between PGM-null 24 and PGM-KO 24 were 25.92 ± 0.41 and 32.21 ± 0.44, respectively. Similarly, the soluble sugars content of PGM-null 38 and PGM-KO 38 were 29.20 ± 0.53 and 35.99 ± 0.45, respectively. Therefore, the result shows that the transgenic effect had a larger impact than the increased carbon source presence in the media, as soluble sugars content of PGM-KO was 75 % higher than PGM-null for both media cases, shown in Figure 4F.

Protein content in dry weight (Dwt) percentage was analyzed for PGM-KO and PGM-null cultures (Figure 4C). The protein contents for both PGM cultures were analyzed based upon elemental analysis. In general, the protein content of PGM-KO embryos was higher at a low C: N mole ratio 24 versus 38, with 4% and 3.3% when compared with PGM-null culture, where Δ = PGM-KO – PGM-null. However, there

is no significant difference in the protein content of PGM-KO 24 and PGM-null 24. To investigate whether the oil content changed due to different amount of sucrose concentrations in the media, hexane extraction was performed on both PGM cultures (Figure 4D). On average, the range of oil content, based on dry weight (Dwt), were 7.07 ± 0.07 and 8.05 ± 0.32 % for PGM-null and PGM-KO embryos, respectively. Therefore, the content of oil has no correlation to the amount of sucrose in the SHaM media, the same trend observed for protein content. Nevertheless, oil content of PGM-KO displayed a 1 Dwt % higher compared with PGM-null embryo for both media conditions tested. To complete the biomass composition balance, the residual biomass fraction was estimated from the mass balance for both PGM cultures. The percentage residual biomass fraction was performed by subtraction of the sum of the biomass composition (protein + oil + starch + soluble sugar + ash + DNA/RNA). Ash content was examined by Thermo-Gravimetric Analysis (TGA) as described in the Materials and Methods Section. DNA/RNA content was assumed to be 5 percentage of dry weight basis (Stephanopoulos et al., 1998). There is no significant difference statistically between the residual biomass fractions for PGM-KO and PGM-null for two media conditions.

Sucrose and glutamine are the two major carbon and nitrogen sources supplied to SHaM embryos (He *et al.*, 2011; Schmidt *et al.*, 2005). The sucrose and glutamine uptake rates were calculated by measuring the depletion of both substrates from the medium at day 0, day 3, and day 6 of culture (see Materials and Methods). In both cases PGM-KO and PGM-null, the glutamine consumption rates were mirrored the protein biosynthesis as shown in Figure 4G. However, there is no significant

difference in the glutamine consumption rates of PGM-null 24 relative to PGM-KO 38. This similar observation on glutamine uptake rate with protein synthesis was perceived from a different soybean cultivar (unpublished data). Figure 4H indicates that the sucrose consumption rate for PGM-null and PGM-KO, both at C: N mole ratio 24, were not statistically different. However, the data illustrates that the PGM-KO 38 consumed the least sucrose in the media.

As mentioned previously, there is a biological triplicate for each media condition and each genotype. For each biological replicate culture, a protein sample was extracted and hydrolyzed into the corresponding amino acids. The hydrolysis procedure was followed by individual amino acid HPLC analysis; sixteen proteinogenic amino acids can be analyzed by using pre-column derivatization with 6-Aminoquinolyl-N-Hydroxysuccinimidyl carbamate (Cohen, 2000). Figure 5 displays the mole percentage proteinogenic amino acid profiles for both PGM cultures with two media conditions. The total protein content decreased with increased C: N mole ratios. However, the overall proteinogenic amino acids did not deviate, with the exception of Arginine, Figure 5. Therefore, the results illustrate that the total protein content respond to the reduced nitrogen source, while the relative proteinogenic amino acid level was not dramatically different.

3.4 ^{13}C-Metabolic Flux Analysis and Steady State Verification of PGM cultures

PGM-KO and PGM-null embryos were cultured in the presence of 10% U-^{13}C-labeled sucrose and the biomass composition mentioned above, and quantification of mass isotopomer fractions in the protein hydrolysate (124), starch hydrolysate (18), and glucosyl units (18) by 2D [^{13}C, ^{1}H] HSQC NMR are shown in Supplemental

Table S1. A mathematical framework of ^{13}C Metabolic Flux Analysis (^{13}C-MFA), NMR2Flux, for metabolic network used here is based on the former studies of developing soybean embryos, and has successfully explained various aspects of central carbon metabolism (Iyer *et al.*, 2008; Sriram *et al.*, 2004). It is a computational approach that mathematically balances the carbon through metabolite pools in a given stoichiometric network. The stoichiometric network accounts for carbon rearrangements and the carbon catalyzed by enzyme reactions, then incorporates isotope label redistribution through different metabolic pathways. Through a statistical fitting process (χ^2), the computational values of fluxes were optimized to account for both the experimental measurement of isotopomer label and the extracellular measurement of flux data, resulting in "best fitted" global flux values. The isotopomer labeling experiments, through the mathematical metabolite model, provided a quantitative description of the carbon flow through biochemical pathways. The metabolic network represents three cellular compartments: cytosol, plastid, and mitochondrion. Labeling data available suggest that the glucosyl units of the protein and starch hydrolysates provided evidence for parallel pathways of glycolysis and pentose phosphate pathways in both cytosol and plastid compartment (Sriram *et al.*, 2007). Further, the mitochondrion composed of the tricarboxylic acid (TCA) cycle, malate shuttle, and anaplerotic reactions were included in the metabolic network model. Glutamine assimilation reactions are considered to occur in the plastid. Recent studies (Allen and Young, 2013; Lonien and Schwender, 2009; Masakapalli *et al.*, 2009; Williams *et al.*, 2008) support the compartmentalization proposed in the soybean model by Sriram et al., (2004). In addition to a previous

developing soybean embryo model, in this study, another variable was added: the fixation of evolved CO_2 by the ribulose 1,5-bisphosphate carboxylase/oxygenase (Rubisco) bypass flux that was suggested by Schwender et al., (2004).

The major assumption of the ^{13}C-MFA model is in the isotopic steady state conditions of the labeling studies of soybean PGM embryos. The steady state flux analysis implies that intracellular metabolite pools do not change over the span of time during which the labeling experiment takes place, which can be achieved after multiple residence times of the metabolite pools. Three measurements can be obtained to validate the metabolic steady state condition: sucrose consumption rates, protein and starch accumulation, and growth characteristic (Kruger *et al.*, 2012). The developing linseed embryos study indicates that the isotopic steady state was about 18 hours (Troufflard *et al.*, 2007). Both PGM embryos were cultured in the sucrose labeled liquid media for 6 days, which shows that the isotopic steady state is attained within eight residence times. Therefore, for both embryos the metabolite pools are expected to be constant, and at steady-state. Furthermore, the core of this study is based on the analysis of the ^{13}C-experiment label in storage products of protein and starch. The products of protein and starch hydrolysate extractions are biosynthesis-derived from the central metabolites. However, for PGM-KO embryos, the amount of starch that was extracted for the hydrolysate and NMR analysis is not significant. As a result, the labeling in proteinogenic amino acids and glycosyl units represents the labeling fingerprint of the central metabolites during protein biosynthesis.

Normalization of the relative flux method (normalization of 100 moles of sucrose uptake) provides a way to elucidate carbon regulation within the metabolic network.

Furthermore, examining relative fluxes and flux ratios of various nodes in the reaction network can enhance the understanding of the metabolic effect on SHaM embryos of null and PGM cultures. As they are genetically different, this understanding could lead to an understanding of these embryos without the influence of the carbon source (Stephanopoulos and Vallino, 1991). From this point on, all references to flux and flux ratios are in terms of relative flux values for the comparison: 1) influence of carbon source in the media for PGM-null and PGM-KO; 2) the transgenic effect of low and high sucrose in the media.

3.5 Metabolic flux comparison of PGM cultures

A labeling experiment using Uniformly-^{13}C Sucrose was used to explore central carbon metabolism under two different levels of carbon supply to soybean PGM-null and PGM-KO embryos. Measurements used for modeling were taken from labeling deposition onto amino acids from hydrolyzed protein and glucosyl units from hydrolyzed protein and starch (Sriram *et al.*, 2004). For both the PGM-null and PGM-KO cultures, protein was hydrolyzed for isotopomer measurement, and the flux map was developed for each of the biological replicates in PGM-null 24, PGM-null 38, PGM-KO 24 and PGM-KO 38 (total 12 flux maps; three biological replicates). After each biological replicate was computed and flux maps were generated for the 12 flux maps in this study, the three biological replicates of each PGM culture condition were averaged and the variance was pooled to incorporate the simulation errors with the biological errors. Therefore, Table 2 displays the average relative flux of each of the PGM culture and the standard deviation was calculated by pool variance. This

provides a measure of accuracy for the biological replicates in each of the PGM culture conditions.

The isotopomeric compositions of the cytosolic and plastidic triose phosphates were not significantly different, based on comparisons of the multiplet intensities of Ala α, Phe α, and Tyr α. The plant biochemistry of soybean indicates that Phe and Tyr synthesize from plastidic PEP and hence reflect the isotopomer composition of the plastidic triose phosphate. Ala is synthesized in cytosol and plastid (http://pmn.plantcyc.org), and therefore the composition displays a combination of isotopomeric composition in the triose phosphates from both compartments. The two genotypes and two C:N mole ratios illustrate that corresponding isotopomer abundances were similar (data not shown). Thus, the data suggests that the flux values of gap, eno, pyk, in cytosol and plastid compartments were indistinguishable. The flux through Rubisco bypass reaction appears to be insignificant for all of the PGM cultures (Table 2). A comparison of simulated and experimented data indicated no gross errors (Figure 6). The statistical consistency between the computationally simulated and the isotopomer-measured data supported the metabolic network model.

Since PGM-KO is impaired in starch synthesis, for both C:N mole ratios, detailed analysis of the flux balance around carbohydrate synthesis reveals that for both genotypes, the carbon flux of polysaccharide synthesis via cytosolic glucose-6-phosphate (bios 4) is influenced by higher C: N ratios (Figure S1A). However, the flux toward the polysaccharide synthesis through cytosolic pentose-5-phosphate (bios 7) indicates a statistical difference between PGM-null and PGM-KO cultures and increase flux toward higher C: N mole ratio of PGM-KO (Figure S1C). As expected,

the flux toward starch synthesis (bios 5) was significantly different between PGM-null and PGM-KO, where PGM-KO 24 and 38 displays little flux through that pathway (Figure S1B). The next set of reactions that are significantly different between the two PGM cultures and the C: N mole ratios were the cytosolic and plastidic pentose phosphate pathway (Figure S2). In the control condition (PGM-null), the flux through non-oxidative reactions of the OxPPP catalyzed by transketolase (tktA and tktB) and transaldolase (tal) were higher in plastid than in the cytosol (Figure S2). However, the mutant embryos indicate a reverse trend where it is higher in cytosol compared to plastid. Hence, knockout of the plastidic phosphoglucomutase gene of soybean embryos affects the cytosolic tkt and tal fluxes. Nonetheless, both PGM-KO 24 and 38 C: N mole ratios do not affect the plastic tkt and tal fluxes. These fluxes are more sensitive to the control culture of PGM-null at different C: N mole ratio of 24 and 38 media conditions (Figure S2). Figure 7 and Table 2 display the fluxes that varied significantly with C: N mole ratio, and are affected by gene knockout. Fluxes are color-coded based on the influence of gene, carbon treatment, and/or combination of both carbon and gene affect (Figure 7).

The flux through cytosolic glucose-6-phosphate isomerase (hxi_f^c) appears to be in a reverse direction of the carbon flow. In four cases, the direction of flux is from fructose-6-phosphate to glucose-6-phosphate. Furthermore, the flux is reduced in the control embryos relative to the gene knockout by approximately 2.5 fold (Figure S3A). However, for both PGM-KO and PGM-null cultures, there is no statistical difference between the two C: N mole ratios of 24 and 38. Likewise, the flux in plastidic fructose-bisphosphate aldolase ($f16bp^p$) displays a similar trend of no

significant difference in the C: N mole ratios of 24 and 38 for both PGM cultures. Nevertheless, the $f16bp^{p}$ fluxes are higher in control culture than in the PGM gene knockout culture (Figure S3B). The plastic triose-phosphate isomerase (pfk^{p}) flux illustrates an increase in the carbon flux as the C: N mole ratios increased in the control culture of PGM-null, but the flux decreased for the PGM knockout cultures (Figure S3C). Therefore, the fluxes of hxi_{f}^{c} and $f16bp^{p}$ are dictated by the transgenic effect of the PGM gene, and the flux through pfk^{p} is influenced by the gene itself, as well as different levels of carbon in the media.

For all cases, there is a significant flux through glycolysis from hexose to total pyruvate, highlighting that sugars are an important source of carbon for fatty acid synthesis. The flux of total pyruvate is shown to be higher in PGM knockout culture relative to the control; however, there is no significant difference between the two C: N mole ratio conditions (Figure S3D). Furthermore, as the PGM culture is influenced by protein and oil biosynthesis, detailed analysis of the mass balance around plastidic pyruvate (Pyr^{p}) indicates the most important features of the flux gene are carbon treatments. This reaction is a crossroad for carbon partitioning between oil and protein metabolism because carbon is reallocated from fatty acid biosynthesis to pyruvate-derived amino acid production. The flux through fatty acid biosynthesis via plastic acetyl-CoA is revealed to be higher in PGM-KO than in the control culture (Figure S3F). Moreover, the pyruvate-derived amino acid leucine biosynthesis displays a similar trend of higher flux in PGM-KO for both C: N mole ratios of 24 and 38. Therefore, this indicates that carbon partitioning was leveraged through pyruvate for protein and oil biosynthesis in the absence of the starch synthesis.

Glutamine uptake is a critical role in the protein biosynthesis as mentioned previously. In the case of PGM 24 and PGM 38 have higher total protein content as compared to the control of PGM-null 24 and PGM-null 38; therefore, the flux of glutamine uptake in our soybean systems mirror the total protein content (Figure S4A). Furthermore, an increased demand of reduced nitrogen source for protein metabolism revealed an increase in glutamate synthesis reaction (as_f) (Figure S4B), but a decrease in the glutamate degradation reaction (gdh_f^p) (Figure S4C). However, the fluxes through the biosynthesis of proline, arginine, proteinogenic glutamine, valine, leucine, alanine, ileucine, serine, glycine, and lysine indicate that the PGM-KO at C: N mole ratio of 38 provided the highest flux towards the synthesis of the protein pool (Figure S5). The same observation was observed for the flux through mitochondrion pyruvate dehydrogenase (pdh^m) for the synthesis of Acetyl-CoA into the TCA cycle (Figure S3E). In addition, glutamine influences the anaplerotic reactions of cytosolic phosphoenolpyruvate carboxylate (ppc^c) and malate dehydrogenase (mdh^c), where the PGM-KO cultures uptake more glutamine resulting in less fluxes through ppc^c and mdh^c (Figures S6A and S6B).

The reaction of ppc^c is converted to malate, which will have a dual function: MalT1 transported flux from cytosol into the mitochondrion for TCA cycle and MalT2 transported flux into the plastid for the synthesis of protein and oil for the control culture. The C: N mole ratio influences the MalT1 transporter in PGM-null cultures. However, the PGM-KO cultures indicate that the direction of carbon flow of MalT1 transported fluxes from mitochondrion for TCA cycle into the cytosolic malate (Figure S6C). Furthermore, there is no statistically different for the MalT1

between PGM-KO 24 and PGM-KO 38. The reversibility of MalT1 for all PGM cultures were very high (> 98%) showing that there was a quick exchange of flux between the two pools. Additionally, the flux through mitochondrion fumarase (sdh_f2^m) illustrates lowest flux in the PGM-KO 38 culture (Figure S6D).

4 Discussion

This study was designed to investigate the use of a soy somatic transgenic embryo (PGM-KO and PGM-null) system as a physiological model of seed development, using steady-state ^{13}C Metabolic Flux Analysis (^{13}C MFA), particularly as a model system for understanding fundamental aspects of controlling resource partitioning. PGM-KO and PGM-null embryos were exposed to various sucrose and glutamine concentrations in SHaM media. To our knowledge, this is the first reported study of flux models on soybean transgenic PGM knockout genes relative to a control, using variations of the carbon supply to both PGM cultures. ^{13}C MFA is able to identify compartmentalization in a plant system (Masakapalli *et al.*, 2010; Sriram *et al.*, 2004), as well as the quantification of responses to environmental perturbations (Allen *et al.*, 2009; Allen and Young, 2013; Iyer *et al.*, 2008) and genetic engineering (Alonso *et al.*, 2007; Lonien and Schwender, 2009). The information from our study provides a numerical comparison of relative growth rate, biomass composition, and central metabolite fluxes on both PGM embryos (KO and null) when cultured in C: N mole ratios of 24 and 38. Moreover, our result from metabolic flux analysis of soybean somatic embryos, under the influence of varying transgenic and media compositions, can help us answer fundamental questions about the controlling of resource partitioning for soybean seed. The isotopic steady-state condition of ^{13}C

MFA was verified for both PGM cultures. The isotopic steady state assumption was based on what was obtained from a linseed embryo case study, which indicated that the steady-state time was 18 hours (Troufflard *et al.*, 2007). Both PGM cultures were cultured in labeling carbon source for 6 days. Therefore, the metabolic flux analysis presented in this study analysis can be considered satisfactory, under the major assumption of isotopic steady state.

From previous studies, we have determined the range of sucrose concentrations to be 88 – 234 mM, and 20 – 60 mM for glutamine concentrations, both of which are consistent with reported values for soybean embryos cultured *in vitro* studies (Allen and Young, 2013; Gifford and Thorne, 1985; Hsu *et al.*, 1984; Pipolo *et al.*, 2004; Saravitz and Raper, 1995; Schmidt *et al.*, 2005). As postulated, under excess carbon and nitrogen in the liquid media, the relative growth rates for both PGM cultures were constant across the broad spectrum of C: N mole ratios from 8.8 to 70.2 during 6-day culture, indicating the presence of an optimal supply of carbon and nitrogen sources for mature soybean somatic embryos to grow (Figure 1). The relative growth rate for PGM-null embryos and PGM-KO embryos were not statistically different, at rates of 0.62 ± 0.08 (day^{-1}) and 0.60 ± 0.08 (day^{-1}), respectively, across the C: N mole ratio of 8.8 to 70.2. For the labeling experiment, the relative growth rates were also not significantly different for PGM-null embryos and PGM-KO embryos, 0.61 ± 0.01 (day-1) and 0.62 ± 0.01 (day^{-1}), respectively, under the C: N mole ratios of 24 and 38 with 10% U-$^{13}C_{12}$ (Table 1). Furthermore, the relative growth rates for both PGM cultures were comparable to those of three different zygotic embryo cultivars: *cv. Evans* with a value of 0.45 ± 0.07 (day^{-1}) (cultured *in* vitro146 mM sucrose with 10%

U-$^{13}C_{12}$, 37 mM glutamine) (Iyer *et al.*, 2008), *cv. Amsoy* with a value of 0.43 ± 0.04 (day^{-1}) (cultured *in vitro* 140 mM sucrose, 70 mM U-$^{13}C_6$, 35 mM U-$^{13}C_5$, 12.5 mM Asparagine) (Allen *et al.*, 2009), and *cv. Jack* (cultured *in vitro* at C:N ratios from 6 to 91 of nearly 40 individual cultures) with an average value of 0.67 ± 0.10 (Day^{-1}) (Allen and Young, 2013). Therefore, observations on the flux distribution in both PGM cultures should be relevant for the development of zygotic embryos. Biomass composition and transcript profiling data of both PGM cultures (data not shown) represent zygotic embryos in the early to mid-storage phase of development.

The soybean somatic embryos with PGM knockout resulted in an insignificant amount of starch in the embryos across the range of C: N mole ratio from 8.8 to 70.2 (Figure 4C). Our current understanding of the classical pathway for starch synthesis in the plastidic compartment is that the first committed step is the conversion of glucose-6-phosphate (G6P) to glucose-1-phosphate (G1P), a reaction catalyzed by phosphoglucomutase. Phosphoglucomutase (PGM) is therefore needed for starch synthesis. Since PGM-KO embryos cannot synthesize starch via the plastid pathway, the synthesis of sucrose from the cytosolic pathway could be increased, as suggested by the soluble sugars data, shown in Figure 4D. The metabolic flux analysis reveals that the flux toward starch synthesis was insignificant in PGM-KO relative to PGM-null, as expected, and the flux via polysaccharide synthesis in cytosol was significantly higher for PGM-null compare to PGM-KO. Elimination of the starch pool provides a way to channel carbon away from the plastidic compartment, and into the cytosolic. The result indicates that the flux through non-oxidative reactions of the cytosolic pentose phosphate pathway (tkt and tal) was statistically higher in PGM-KO

than in PGM-null (Figure 7), and it was due to higher flux of pentose sugars synthesis in the system (bios7). The soluble sugar pool was mainly synthesized from non-oxidative transketolase PPP reaction, by transferring a keto group to glyceraldehydes-3P from fructose-6P. As a result, the cytosolic glucose-6-phosphate isomerase (hxi_f^c) flux was significantly higher in PGM-KO than in PGM-null. Therefore, manipulation of starch synthesis in the system heavily affects the cytosolic non-oxidative pentose phosphate pathway.

The deletion of PGM (PGM-KO) has a positive effect on protein and oil synthesis. PGM-KO has consistently greater increases in both protein and oil when compared to PGM-null, for protein (Δ2.3 Dwt % (average); Δ4 Dwt % (maximum)) and oil (Δ1.8 Dwt % (average); Δ5.5 Dwt % (maximum)) across the range of C: N mole ratios, where Δ = PGM-KO – PGM-null. As illustrated in Figure 3, PGM-null and PGM-KO embryos utilize the glutamine from the media as a primary source of nitrogen for storage protein biosynthesis. The total protein content correlated strongly with C:N ratios in the media for PGM-null and PGM-KO respectively, at $R^2 = 0.79$ and $R^2 = 0.80$ (data not shown). The total protein content also correlated strongly with glutamine consumption rate, at $R^2 = 0.74$ and $R^2 = 0.82$ for PGM-null and PGM-KO, respectively (Figure 3). These results indicate that protein biosynthesis in soybean somatic PGM embryos depends on reduced nitrogen source in the liquid media. Thus, in the absence of PGM, the system uptake more glutamine, and demonstrated higher glutamate metabolism for the role of higher protein synthesis (Figure 7). As expected, amino acid biosynthesis for PGM-KO was statistically higher than PGM-null. This is consistent with the recent result, from *in vitro* soybean zygotic embryos, which

indicated that glutamine provides approximately 32% to 46% carbon toward amino acid synthesis (Allen and Young, 2013). Moreover, the nitrogen source was well known for its role in the protein biosynthesis of legume (Haga and Sodek, 1987; Thompson *et al.*, 1977). The conversion of glutamine to the other amino acids that are required for storage protein synthesis demands carbon skeletons derived from metabolism of sucrose to be imported from the phloem (Miflin and Lea, 1977). In the case of glutamine, its amide-N group must be donated to a suitable C-skeleton acceptor molecule, such as 2-Oxoglutarate (a product of TCA cycle in the mitochondria), which can then be utilized in the formation of other amino acids (Weigelt et al., 2008).

Furthermore, the high rate of glutamine uptake with a constant carbon source in both PGM culture could possibly influence anaplerotic reactions. The anaplerotic reactions catalyzed by cytosolic ppp^{c}, me^{p}, and malate (*MalT1* and *MalT2*) transporters for shunting malate between cytosol, plastid, and mitochondrion were investigated. The flux through the anaplerotic reaction of ppc^{c} was highest in PGM-null 38, as shown in Figure S6A. Flux analysis of PGM-KO culture indicates that the malate transporter (*MalT1*) preferred to channel carbon from the mitochondrion TCA cycle into the cytosolic malate for protein and oil synthesis. Hence, the oil and protein fluxes were statistically higher in PGM-KO as compare to PGM-null. The malate transporter was found to be responsible for the carbon partitioning of protein and oil content from the plastidic pyruvate pool in a metabolic flux analysis of soybean zygotic embryos with various temperatures (Iyer *et al.*, 2008). Furthermore, the result agrees with the recent metabolic flux analysis of soybean zygotic *in vitro* studies by

revealing that 9% to 19% of carbon toward fatty acid biosynthesis came from the supply of glutamine (Allen and Young, 2013). Moreover, for both PGM cultures, malic enzymes provide a small fraction of pyruvate entering the TCA cycle, which is consistent with previous soy zygotic embryos studies (Allen *et al.*, 2009; Allen and Young, 2013; Iyer *et al.*, 2008). Therefore, the fluxes through the TCA cycle and those enzymes could be potential targets of metabolic engineering for improving protein biosynthesis within the soybean developing embryos.

Acknowledgements

The authors would like to thank undergraduate student, Miya Williams, for extracted 2D NMR spectra. We would like to acknowledge Dr. D. Bruce Fulton (Department of Biochemistry, Biophysics and Molecular Biology, Iowa State University) for help with the NMR experiments. This work was supported by DuPont Pioneer Hi-Bred.

References

Agrawal GK, Hajduch M, Graham K, Thelen JJ. 2008. In-depth investigation of the soybean seed-filling proteome and comparison with a parallel study of rapeseed. *Plant Physiology* 148, 504-518.

Allen DK, Libourel IGL, Shachar-Hill Y. 2009. Metabolic flux analysis in plants: coping with complexity. *Plant Cell and Environment* 32, 1241-1257.

Allen DK, Ohlrogge JB, Shachar-Hill Y. 2009. The role of light in soybean seed filling metabolism. *Plant Journal* 58, 220-234.

Allen DK, Shachar-Hill Y, Ohlrogge JB. 2007. Compartment-specific labeling information in C-13 metabolic flux analysis of plants. *Phytochemistry* 68, 2197-2210.

Allen DK, Young JD. 2013. Carbon and Nitrogen Provisions Alter the Metabolic Flux in Developing Soybean Embryos. *Plant Physiology* 161, 1458-1475.

Allen DK, Young JD. 2013. Carbon and Nitrogen Provisions Alter the Metabolic Flux in Developing Soybean Embryos. *Plant Physiology*.

Allen SM, Butler KH, Carlson TJ, Hitz WD, Stoop JM. 2008. Plastidic phosphoglucomutase genes. Vol. US 20080066204A1: EI du Pont de Nemours and Company.

Alonso AP, Dale VL, Shachar-Hill Y. 2010. Understanding fatty acid synthesis in developing maize embryos using metabolic flux analysis. *Metabolic Engineering* 12, 488-497.

Alonso AP, Goffman FD, Ohlrogge JB, Shachar-Hill Y. 2007. Carbon conversion efficiency and central metabolic fluxes in developing sunflower (Helianthus annuus L.) embryos. *Plant Journal* 52, 296-308.

Alonso AP, Piasecki RJ, Wang Y, LaClair RW, Shachar-Hill Y. 2010. Quantifying the Labeling and the Levels of Plant Cell Wall Precursors Using Ion Chromatography Tandem Mass Spectrometry. *Plant Physiology* 153, 915-924.

Alonso AP, Raymond P, Hernould M, et al. 2007. A metabolic flux analysis to study the role of sucrose synthase in the regulation of the carbon partitioning in central metabolism in maize root tips. *Metabolic Engineering* 9, 419-432.

Alonso AP, Raymond P, Rolin D, Dieuaide-Noubhani M. 2007. Substrate cycles in the central metabolism of maize root tips under hypoxia. *Phytochemistry* 68, 2222-2231.

Alonso AP, Val DL, Shachar-Hill Y. 2011. Central metabolic fluxes in the endosperm of developing maize seeds and their implications for metabolic engineering. *Metabolic Engineering* 13, 96-107.

Alonso AP, Val DL, Shachar-Hill Y. 2011. Understanding fatty acid synthesis in developing maize embryos using metabolic flux analysis (vol 12, pg 488, 2010). *Metabolic Engineering* 13, 454-454.

Bailey MA, Boerma HR, Parrott WA. 1993. Genotype-Specific Optimization of Plant-Regeneration from Somatic Embryos of Soybean. *Plant Science* 93, 117-120.

Baud S, Lepiniec L. 2010. Physiological and developmental regulation of seed oil production. *Progress in Lipid Research* 49, 235-249.

Bodenhausen G, Ruben DJ. 1980. Natural abundance nitrogen-15 NMR by enhanced heteronuclear spectroscopy. *Chemical Physics Letters* 69, 185-189.

Boghigian BA, Seth G, Kiss R, Pfeifer BA. 2010. Metabolic flux analysis and pharmaceutical production. *Metabolic Engineering* 12, 81-95.

Borisjuk L, Nguyen TH, Neuberger T, et al. 2005. Gradients of lipid storage, photosynthesis and plastid differentiation in developing soybean seeds. *New Phytologist* 167, 761-776.

Bowsher CG, Tobin AK. 2001. Compartmentation of metabolism within mitochondria and plastids. *Journal of Experimental Botany* 52, 513-527.

Boyle NR, Morgan JA. 2009. Flux balance analysis of primary metabolism in Chlamydomonas reinhardtii. *BMC Systems Biology* 3, (7 January 2009).

Bradford MM. 1976. Rapid and Sensitive Method for Quantitation of Microgram Quantities of Protein Utilizing Principle of Protein-Dye Binding. *Analytical Biochemistry* 72, 248-254.

Brown LR. 1984. Differential Scaling along Omega-1 in COSY Experiments *Journal of Magnetic Resonance* 57, 513-518.

Carrari F, Urbanczyk-Wochniak E, Willmitzer L, Fernie AR. 2003. Engineering central metabolism in crop species: learning the system. *Metabolic Engineering* 5, 191-200.

Chanprame S, Kuo TM, Widholm JM. 1998. Soluble Carbohydrate Content of Soybean [Glycine max (L.) Merr.] Somatic and Zygotic Embryos during Development. *In Vitro Cellular & Developmental Biology. Plant* 34, 64-68.

Clemente TE, Cahoon EB. 2009. Soybean Oil: Genetic Approaches for Modification of Functionality and Total Content. *Plant Physiology* 151, 1030-1040.

Cohen SA. 2000. Amino Acid Analysis Using Precolumn Derivatization with 6-Aminoquinolyl-N-Hydroxysuccinimidyl Carbamate. *Methods in Molecular Biology: Amino Acid Analysis Protocols* 159, 39-47.

Dietz K. 1987. Control function of hexosemonophosphate isomerase and phosphoglucomutase in starch synthesis in leaves. *In J Biggins, ed, Proceedings of the 7th International Congress on Photosynthesis* 3, 329-332.

Dieuaide-Noubhani M, Alonso AP, Rolin D, Eisenreich W, Raymond P. 2007. *Metabolic flux analysis: recent advances in carbon metabolism in plants*.

Durot M, Bourguignon PY, Schachter V. 2009. Genome-scale models of bacterial metabolism: reconstruction and applications. *FEMS Microbiology Reviews* 33, 164-190.

Fait A, Fromm H, Walter D, Galili G, Fernie AR. 2008. Highway or byway: the metabolic role of the GABA shunt in plants. *Trends in Plant Science* 13, 14-19.

Feist AM, Herrgard MJ, Thiele I, Reed JL, Palsson BO. 2009. Reconstruction of biochemical networks in microorganisms. *Nature Reviews Microbiology* 7, 129-143.

Finer JJ, McMullen MD. 1991. Transformation of Soybean Via Particle Bombardment of Embryogenic Suspension Culture Tissue. *In Vitro Cellular and Developmental Biology Plant* 27P, 175-182.

Gifford RM, Thorne JH. 1985. Sucrose Concentration at the Apoplastic Interface Between Seed Coat and Cotyledons of Developing Soybean Seeds. *Plant Physiology* 77, 863-868.

Golombek S, Rolletschek H, Wobus U, Weber H. 2001. Control of storage protein accumulation during legume seed development. *Journal of Plant Physiology* 158, 457-464.

Gonzalez DO, Vodkin LO. 2007. Specific elements of the glyoxylate pathway play a significant role in the functional transition of the soybean cotyledon during seedling development. *Bmc Genomics* 8.

Haga KI, Sodek L. 1987. Utilization of Nitrogen-Sources by Immature Soybean Cotyledons in Culture. *Annals of Botany* 59, 597-601.

Harris RK. 1983. Nuclear Magnetic Resonance Spectroscopy: A Physiochemical View. *Pitman Books, London*.

Hartwig EE, Kilen TC. 1991. Yield and Composition of Soybean Seed from Parents with Different Protein, Similar Yield. *Crop Science* 31, 290-292.

Hattenbach A, Heineke D. 1999. On the role of chloroplastic phosphoglucomutase in the regulation of starch turnover. *Planta* 207, 527-532.

Hay J, Schwender J. 2011. Computational analysis of storage synthesis in developing Brassica napus L. (oilseed rape) embryos: flux variability analysis in relation to (13)C metabolic flux analysis. *Plant Journal* 67, 513-525.

Hay J, Schwender J. 2011. Metabolic network reconstruction and flux variability analysis of storage synthesis in developing oilseed rape (Brassica napus L.) embryos. *Plant Journal* 67, 526-541.

Hayati R, Egli DB, CraftsBrandner SJ. 1996. Independence of nitrogen supply and seed growth in soybean: Studies using an in vitro culture system. *Journal of Experimental Botany* 47, 33-40.

He Y, Young T, Clark K, Kleppinger-Sparace K, Bridges W, Sparace S. 2011. Developmental profile of storage reserve accumulation in soybean somatic embryos. *In Vitro Cellular & Developmental Biology - Plant* 47, 725-733.

Hernandez-Sebastia C, Marsolais F, Saravitz C, Israel D, Dewey RE, Huber SC. 2005. Free amino acid profiles suggest a possible role for asparagine in the control of storage-product accumulation in developing seeds of low- and high-protein soybean lines. *Journal of Experimental Botany* 56, 1951-1963.

Hsu FC, Bennett AB, Spanswick RM. 1984. Concentrations of Sucrose and Nitrogenous Compounds in the Apoplast of Developing Soybean Seed Coats and Embryos *Plant Physiology* 75, 181-186.

Hymowitz T, Collins FI, Walker WM, Panczner J. 1972. Relationship between content of oil, protein, and sugar in soybean seed. *Agronomy Journal* 64, 613-616.

Iyer VV, Sriram G, Fulton DB, Zhou R, Westgate ME, Shanks JV. 2008. Metabolic flux maps comparing the effect of temperature on protein and oil biosynthesis in developing soybean cotyledons. *Plant Cell and Environment* 31, 506-517.

Iyer VV, Sriram G, Shanks JV. 2007. Metabolic flux maps of central carbon metabolism in plant systems - Isotope labeling analysis. *Concepts in Plant Metabolomics*, 125-144.

Johnson BA, Blevins RA. 1994. NMR View - A Computure-Program for the Visualization and Analysis of NMR data. *Journal of Biomolecular Nmr* 4, 603-614.

Joshi V, Laubengayer KM, Schauer N, Fernie AR, Jander G. 2006. Two Arabidopsis threonine aldolases are nonredundant and compete with threonine deaminase for a common substrate pool. *Plant Cell* 18, 3564-3575.

Junker BH, Lonien J, Heady LE, Rogers A, Schwender J. 2007. Parallel determination of enzyme activities and in vivo fluxes in Brassica napus embryos grown on organic or inorganic nitrogen source. *Phytochemistry* 68, 2232-2242.

Kell DB. 2004. Metabolomics and systems biology: making sense of the soup. *Current Opinion in Microbiology* 7, 296-307.

Kinney AJ. 1996. Development of genetically engineered soybean oils for food applications. *Journal of Food Lipids* 3, 273-292.

Koch K. 2004. Sucrose metabolism: regulatory mechanisms and pivotal roles in sugar sensing and plant development. *Current Opinion in Plant Biology* 7, 235-246.

Koch KE. 1996. Carbohydrate-modulated gene expression in plants. *Annual Review of Plant Physiology and Plant Molecular Biology* 47, 509-540.

Krivdin LB, Kalabin GA. 1989. Structural Applications of One-Bond Carbon-Carbon Spin-Spin Coupling-Constants *Progress in Nuclear Magnetic Resonance Spectroscopy* 21, 293-448.

Kruger NJ, Masakapalli SK, Ratcliffe RG. 2012. Strategies for investigating the plant metabolic network with steady-state metabolic flux analysis: lessons from an Arabidopsis cell culture and other systems. *Journal of Experimental Botany* 63, 2309-2323.

Kruger NJ, Ratcliffe RG. 2009. Insights into plant metabolic networks from steady-state metabolic flux analysis. *Biochimie* 91, 697-702.

Kruger NJ, Troncoso-Ponce MA, Ratcliffe RG. 2008. 1H NMR metabolite fingerprinting and metabolomic analysis of perchloric acid extracts from plant tissues. *Nat. Protocols* 3, 1001-1012.

Kruger NJ, von Schaewen A. 2003. The oxidative pentose phosphate pathway: structure and organisation. *Current Opinion in Plant Biology* 6, 236-246.

Last RL, Jones AD, Shachar-Hill Y. 2007. Towards the plant metabolome and beyond. *Nature Reviews Molecular Cell Biology* 8, 167-174.

Li ZS, Moon BP, Xing AQ, Liu ZB, McCardell RP, Damude HG, Falco SC. 2010. Stacking Multiple Transgenes at a Selected Genomic Site via Repeated Recombinase-Mediated DNA Cassette Exchanges. *Plant Physiology* 154, 622-631.

Libourel IGL, Shachar-Hill Y. 2008. Metabolic flux analysis in plants: from intelligent design to rational engineering. *Annual Review of Plant Biology* 59, 625-650.

Litterer LA, Plaisance KL, Schnurr JA, Storey KK, Jung HJG, Gronwald JW, Somers DA. 2006. Biosynthesis of UDP-glucuronic acid in developing soybean embryos: possible role of UDP-sugar pyrophosphorylase. *Physiologia Plantarum* 128, 200-211.

Lonien J, Schwender J. 2009. Analysis of Metabolic Flux Phenotypes for Two Arabidopsis Mutants with Severe Impairment in Seed Storage Lipid Synthesis. *Plant Physiology* 151, 1617-1634.

Lu C, Napier JA, Clemente TE, Cahoon EB. 2011. New frontiers in oilseed biotechnology: meeting the global demand for vegetable oils for food, feed, biofuel, and industrial applications. *Current Opinion in Biotechnology* 22, 252-259.

Manjunath S, Lee CHK, VanWinkle P, Bailey-Serres J. 1998. Molecular and biochemical characterization of cytosolic phosphoglucomutase in maize - Expression during development and in response to oxygen deprivation. *Plant Physiology* 117, 997-1006.

Masakapalli SK, Lay PL, Huddleston JE, Pollock NL, Kruger NJ, Ratcliffe RG. 2009. Subcellular flux analysis of central metabolism in a heterotrophic Arabidopsis thaliana cell suspension using steady-state stable isotope labeling. *Plant Physiology*, pp.109.151316.

Masakapalli SK, Le Lay P, Huddleston JE, Pollock NL, Kruger NJ, Ratcliffe RG. 2010. Subcellular Flux Analysis of Central Metabolism in a Heterotrophic Arabidopsis Cell Suspension Using Steady-State Stable Isotope Labeling. *Plant Physiology* 152, 602-619.

Massou S, Nicolas C, Letisse F, Portais JC. 2007. NMR-based fluxomics: Quantitative 2D NMR methods for isotopomers analysis. *Phytochemistry* 68, 2330-2340.

Medini D, Serruto D, Parkhill J, Relman DA, Donati C, Moxon R, Falkow S, Rappuoli R. 2008. Microbiology in the post-genomic era. *Nature Reviews Microbiology* 6, 419-430.

Meyer K, Hitz WD, Yadav NS, Damude HG. 2012. DGAT genes from Yarrowia lipolytica for increased seed storage lipid production and altered fatty acid profiles in soybean. E I du Pont de Nemours and Company. US Patent No US 8143473.

Meyer K, Kinney AJ. 2009. Biosynthesis and Biotechnology of Seed Lipids Including Sterols, Carotenoids and Tocochromanols. *Lipids in Photosynthesis: Essential and Regulatory Functions*, 407-444,CP414,CP416.

Meyer K, Stecca KL, Ewell-Hicks K, Allen SM, Everard JD. 2012. Oil and Protein Accumulation in Developing Seeds Is Influenced by the Expression of a Cytosolic Pyrophosphatase in Arabidopsis. *Plant Physiology* 159, 1221-1234.

Miflin BJ, Lea PJ. 1977. Amino- Acid Metabolism *Briggs, Winslow R. (Ed.). Annual Review of Plant Physiology, Vol. 28. X+615p. Illus. Annual Reviews Inc.: Palo Alto, Calif., USA. Isbn 0-8243-0628-7*, 299-329.

Miflin BJ, Lea PJ. 1977. Amino Acid Metabolism. *Annual Review of Plant Physiology* 28, 299-329.

Miranda M, Borisjuk L, Tewes A, Heim U, Sauer N, Wobus U, Weber H. 2001. Amino acid permeases in developing seeds of Vicia faba L.: expression precedes storage protein synthesis and is regulated by amino acid supply. *Plant Journal* 28, 61-71.

Morgan JA, Barney CS, Penn AH, Shanks JV. 2000. Effects of buffered media upon growth and alkaloid production of Catharanthus roseus hairy roots. *Applied Microbiology and Biotechnology* 53, 262-265.

Nargund S, Sriram G. 2013. Designer labels for plant metabolism: statistical design of isotope labeling experiments for improved quantification of flux in complex plant metabolic networks. *Molecular Biosystems* 9, 99-112.

Nielsen J, Oliver S. 2005. The next wave in metabolome analysis. *Trends in Biotechnology* 23, 544-546.

Nishizawa K, Ishimoto M. 2009. Maturation of somatic embryos as a model for soybean seed development. *Plant Biotechnology* 26, 543-550.

Nishizawa K, Takagi K, Teraishi M, Kita A, Ishimoto M. 2010. Application of somatic embryos to rapid and reliable analysis of soybean seed components by RNA interference-mediated gene silencing. *Plant Biotechnology* 27, 409-420.

Nishizawaa K, Ishimoto M. 2009. Maturation of somatic embryos as a model for soybean seed development. *Plant Biotechnology* 26, 543-550.

Norton G, Harris JF. 1975. Compositional Changes in Developing Rape Seed (Brassica-Napus L) *Planta* 123, 163-174.

O'Grady J, Schwender J, Shachar-Hill Y, Morgan JA. 2012. Metabolic cartography: experimental quantification of metabolic fluxes from isotopic labelling studies. *Journal of Experimental Botany* 63, 2293-2308.

Obendorf RL, Rytko GT, Byrne MC. 1982. Soya Bean Seed Growth and Maturation by In Vitro Pod Culture. *Ann Bot* 51, 217-227.

Obendorf RL, Rytko GT, Byrne MC. 1983. Soya Bean Seed Growth and Maturation by In Vitro Pod Culture. *Ann Bot* 51, 217-227.

Pandurangan S, Pajak A, Molnar SJ, et al. 2012. Relationship between asparagine metabolism and protein concentration in soybean seed. *Journal of Experimental Botany* 63, 3173-3184.

Paula Alonso A, Dale VL, Shachar-Hill Y. 2010. Understanding fatty acid synthesis in developing maize embryos using metabolic flux analysis. *Metabolic Engineering* 12, 488-497.

Periappuram C, Steinhauer L, Barton DL, Taylor DC, Chatson B, Zou J. 2000. The plastidic phosphoglucomutase from Arabidopsis. A reversible enzyme reaction with an important role in metabolic control. *Plant Physiology* 123, 1197-1197.

Pipolo AE, Sinclair TR, Camara GMS. 2004. Protein and oil concentration of soybean seed cultured in vitro using nutrient solutions of differing glutamine concentration. *Annals of Applied Biology* 144, 223-227.

Ratcliffe RG, Shachar-Hill Y. 2001. Probing plant metabolism with NMR. *Annual Review of Plant Physiology and Plant Molecular Biology* 52, 499-526.

Ratcliffe RG, Shachar-Hill Y. 2006. Measuring multiple fluxes through plant metabolic networks. *Plant Journal* 45, 490-511.

Rolletschek H, Radchuk R, Klukas C, Schreiber F, Wobus U, Borisjuk L. 2005. Evidence of a key role for photosynthetic oxygen release in oil storage in developing soybean seeds. *New Phytologist* 167, 777-786.

Rontein D, Dieuaide-Noubhani M, Dufourc EJ, Raymond P, Rolin D. 2002. The metabolic architecture of plant cells - Stability of central metabolism and flexibility of anabolic pathways during the growth cycle of tomato cells. *Journal of Biological Chemistry* 277, 43948-43960.

Ruuska SA, Schwender J, Ohlrogge JB. 2004. The capacity of green oilseeds to utilize photosynthesis to drive biosynthetic processes. *Plant Physiology* 136, 2700-2709.

Salon C, Munier-Jolain NG, Duc G, Voisin AS, Grandgirard D, Larmure A, Emery RJN, Ney B. 2001. Grain legume seed filling in relation to nitrogen acquisition: A review and prospects with particular reference to pea. *Agronomie* 21, 539-552.

Saravitz CH, Raper CD. 1995. Responses to Sucrose and Glutamine by Soybean Embryos Growth *In-Vitro*. *Physiologia Plantarum* 93, 799-805.

Sauer U. 2006. Metabolic networks in motion: C-13-based flux analysis. *Molecular Systems Biology* 2.

Schmidt MA, Tucker DM, Cahoon EB, Parrott WA. 2005. Towards normalization of soybean somatic embryo maturation. *Plant Cell Reports* 24, 383-391.

Schmutz J, Cannon SB, Schlueter J, et al. 2010. Genome sequence of the palaeopolyploid soybean. *Nature* 463, 178-183.

Schnarrenberger C, Oeser A, Tolbert NE. 1973. Two isoenzymes each of glucose-6-phosphate dehydrogenase and 6-phosphogluconate dehydrogenase in spinach leaves. *Archives of Biochemistry and Biophysics* 154, 438-448.

Schwender J. 2008. Metabolic flux analysis as a tool in metabolic engineering of plants. *Current Opinion in Biotechnology* 19, 131-137.

Schwender J, Goffman F, Ohlrogge JB, Shachar-Hill Y. 2004. Rubisco without the Calvin cycle improves the carbon efficiency of developing green seeds. *Nature* 432, 779-782.

Schwender J, Ohlrogge J, Shachar-Hill Y. 2004. Understanding flux in plant metabolic networks. *Current Opinion in Plant Biology* 7, 309-317.

Schwender J, Ohlrogge JB. 2002. Probing in vivo metabolism by stable isotope labeling of storage lipids and proteins in developing Brassica napus embryos. *Plant Physiology* 130, 347-361.

Schwender J, Ohlrogge JB, Shachar-Hill Y. 2003. A flux model of glycolysis and the oxidative pentosephosphate pathway in developing Brassica napus embryos. *Journal of Biological Chemistry* 278, 29442-29453.

Schwender J, Shachar-Hill Y, Ohlrogge JB. 2006. Mitochondrial metabolism in developing embryos of Brassica napus. *Journal of Biological Chemistry* 281, 34040-34047.

Shachar-Hill Y. 2002. Nuclear magnetic resonance and plant metabolic engineering. *Metabolic Engineering* 4, 90-97.

Sinclair TR, de Wit CT. 1975. Photosynthate and Nitrogen Requirements for Seed Production by Various Crops. *Science* 189, 565-567.

Smeekens S. 2000. Sugar-induced signal transduction in plants. *Annual Review of Plant Physiology and Plant Molecular Biology* 51, 49-81.

Smith AJ, Rinne RW, Seif RD. 1989. Phosphoenolpyruvate Carboxylase and Pyruvate Kinase Involvement in Protein and Oil Biosynthesis during Soybean Seed Development. *Crop Sci.* 29, 349-353.

Sriram G, Fulton DB, Iyer VV, Peterson JM, Zhou R, Westgate ME, Spalding MH, Shanks JV. 2004. Quantification of compartmented metabolic fluxes in developing soybean embryos by employing biosynthetically directed fractional ^{13}C labeling, two-dimensional [^{13}C, ^{1}H] nuclear magnetic resonance, and comprehensive isotopomer balancing. *Plant Physiology* 136, 3043-3057.

Sriram G, Fulton DB, Iyer VV, Peterson JM, Zhou RL, Westgate ME, Spalding MH, Shanks JV. 2004. Quantification of compartmented metabolic fluxes in developing soybean embryos by employing Biosynthetic ally directed fractional C-13 labeling, C-13, H-1 two-dimensional nuclear magnetic resonance, and comprehensive isotopomer balancing. *Plant Physiology* 136, 3043-3057.

Sriram G, Fulton DB, Shanks JV. 2007. Flux quantification in central carbon metabolism of Catharanthus roseus hairy roots by C-13 labeling and comprehensive bondomer balancing. *Phytochemistry* 68, 2243-2257.

Sriram G, Iyer VV, Fulton DB, Shanks JV. 2007. Identification of hexose hydrolysis products in metabolic flux analytes: A case study of levulinic acid in plant protein hydrolysate. *Metabolic Engineering* 9, 442-451.

Stepansky A, Leustek T. 2006. Histidine biosynthesis in plants. *Amino Acids* 30, 127-142.

Stephanopoulos G. 1999. Metabolic Fluxes and Metabolic Engineering. *Metabolic Engineering* 1, 1-11.

Stephanopoulos G. 2002. Metabolic engineering: Perspective of a chemical engineer. *Aiche Journal* 48, 920-926.

Stephanopoulos G, Vallino JJ. 1991. Network Rigidity and Metabolic Engineering in Metabolite Overproduction. *Science* 252, 1675-1681.

Stephanopoulos GN, Aristidou AA, Nielsen J. 1998. Metabolic Flux Analysis. *Metabolic Engineering*. San Diego: Academic Press, 309-351.

Stombaugh SK, Jung HG, Orf JH, Somers DA. 2000. Genotypic and environmental variation in soybean seed cell wall polysaccharides. *Crop Science* 40, 408-412.

Stombaugh SK, Orf JH, Jung HG, Somers DA. 2003. Relationships between soybean seed cell wall polysaccharides, yield, and seed traits. *Crop Science* 43, 571-576.

Sweetlove LJ, Fell D, Fernie AR. 2008. Getting to grips with the plant metabolic network. *Biochem J* 409, 27-41.

Szyperski T. 1995. Biosynthetically directed fractional ^{13}C-labeling of proteinogenic amino acids. An efficient analytical tool to investigate intermediary metabolism. *European Journal of Biochemistry* 232, 433-448.

Szyperski T. 1995. Biosynthetically Directed Fractional C-13-Labeling of Proteinogenic Amino-Acids - An Efficient Analytical Tool to Investigate Intermediary Metabolism. *European Journal of Biochemistry* 232, 433-448.

Thompson JF, Madison JT, Muenster AME. 1977. *In Vitro* Culture of Immature Cotyledons of Soybean (*Glycine-Max-L-Merr*) *Annals of Botany* 41, 29-39.

Troufflard S, Roscher A, Thomasset B, Barbotin JN, Rawsthorne S, Portais JC. 2007. In vivo C-13 NMR determines metabolic fluxes and steady state in, linseed embryos. *Phytochemistry* 68, 2341-2350.

van der Werf MJ, Overkamp KM, Muilwijk B, Coulier L, Hankemeier T. 2007. Microbial metabolomics: Toward a platform with full metabolome coverage. *Analytical Biochemistry* 370, 17-25.

van Winden W, Schipper D, Verheijen P, Heijnen J. 2001. Innovations in generation and analysis of 2D C-13, H-1 COSYNMR spectra for metabolic flux analysis purposes. *Metabolic Engineering* 3, 322-343.

van Winden WA, Heijnen JJ, Verheijen PJT, Grievink J. 2001. A priori analysis of metabolic flux identifiability from C-13-labeling data. *Biotechnology and Bioengineering* 74, 505-516.

Varma A, Palsson BO. 1994. Stoichiometric Flux Balance Models Quantitatively Predict Growth and Metabolic By-Product Secretion in Wild-Type *Escherichia-Coli W3110*. *Applied and Environmental Microbiology* 60, 3724-3731.

Wakao S, Andre C, Benning C. 2008. Functional analyses of cytosolic glucose-6-phosphate dehydrogenases and their contribution to seed oil accumulation in Arabidopsis. *Plant Physiology* 146, 277-288.

Walker DR, Parrott WA. 2001. Effect of polyethylene glycol and sugar alcohols on soybean somatic embryo germination and conversion. *Plant Cell Tissue and Organ Culture* 64, 55-62.

Weigelt K, Kuster H, Radchuk R, Muller M, Weichert H, Fait A, Fernie AR, Saalbach I, Weber H. 2008. Increasing amino acid supply in pea embryos reveals specific interactions of N and C metabolism, and highlights the importance of mitochondrial metabolism. *Plant Journal* 55, 909-926.

Wiechert W. 2001. C-13 metabolic flux analysis. *Metabolic Engineering* 3, 195-206.

Wiechert W, Mollney M, Petersen S, de Graaf AA. 2001. A universal framework for C-13 metabolic flux analysis. *Metabolic Engineering* 3, 265-283.

Wilcox JR, Shibles RM. 2001. Interrelationships among seed quality attributes in soybean. *Crop Science* 41, 11-14.

Williams TCR, Miguet L, Masakapalli SK, Kruger NJ, Sweetlove LJ, Ratcliffe RG. 2008. Metabolic network fluxes in heterotrophic arabidopsis cells: stability of the flux distribution under different oxygenation conditions. *Plant Physiology* 148, 704-718.

Willker W, Flogel U, Leibfritz D. 1997. Ultra-high-resolved HSQC spectra of multiple-C-13-labeled biofluids. *Journal of Magnetic Resonance* 125, 216-219.

Wilson LA. 1995. Soy foods. In D. R. Erickson (ed.) Practial handbook of soybean processing and utilization. AOCS Press, Champaign, IL and United Soybean Board, St. Louis, MO., 428-459.

Wuthrich K. 1976. NMR in Biological Research: Peptides and Proteins. *North Holland, Amsterdam*.

Figures and Legends

Figure 1: The relative growth rates of soybean somatic embryos of PGM-null (A), PGM-KO (B), day-1, with different initial Carbon-to-Nitrogen (C: N) mole ratios. The embryos were cultured in SHaM media for 6 days with a media change on day 3. Error bars represent standard error (n=3). The diamonds show the confidence limits ($\alpha = 0.05$) for each mean (determined by analysis of variance (ANOVA) of the three biological replicates at each media condition).

Figure 2: Soybean somatic embryos of PGM-null, (A) the oil, protein, starch, (B) soluble sugars and residual biomass; PGM-KO, (C) the oil, protein, starch, (D) soluble sugars and residual biomass, contents, all expressed on a dry weight basis, cultured in SHaM media with different initial Carbon-to-Nitrogen (C: N) mole ratios for 6 days. The media was replaced after 3 days in culture. Error bars represent standard error (n = 3).

Figure 3: Liner fit of the glutamine consumption rate of soybean somatic embryos of PGM-null (A), PGM-KO (B), µmol g Dwt^{-1} day^{-1}, versus the protein content expressed on a dry weight basis, cultured in SHaM media with different initial Carbon-to-Nitrogen (C: N) mole ratios for 6 days. Each media condition represents the three biological replicates. The media in each were replaced after 3 days of culture.

Figure 4: Soybean somatic embryos of PGM-null and PGM-KO cultured in SHaM media with carbon labeling experiments of two different C: N mole ratios for 6 days with the media replace after 3 day. (A) Relative growth rate (day-1), (B) starch content (%), (C) protein content (%), (D) oil content (%), (E) residual biomass content (%), (F) soluble sugars content (%), (G) glutamine consumption rate (umol gDwt-1 day-1), (H) sucrose consumption rate (umol gDwt-1 day-1), all expressed on a dry weight basis. Data represent the means of three biological replicates with standard errors. The means were statistically compared within each culture with an analysis of variance followed by student's t mean comparison ($P < 0.05$). Means within each treatment group with the same letter are not significantly different.

Figure 5: The proteinogenic amino acid levels were measured in three biological replicates for soybean somatic embryos of PGM-null and PGM-KO for two of the C: N mole ratios. Embryos were cultured for 6 days and each culture the media was replaced after day 3.

Figure 6: Experimented versus simulated data. The experimentally determined from 2D NMR HSQC isotopomer labeling patterns were compared with the simulated values from metabolic flux analysis of each of the four flux maps. Measurements

were in consistency with model, illustrating that the models were a good fit with the metabolic network topology of the soybean system model.

Figure 7: Metabolic network model of central carbon metabolism in soybean somatic embryos PGM cultures with supplied two different C: N mole ratios, based on the network described earlier for flux studies with developing soybean zygotic embryos (Sriram et al., 2004). Color code in reach indicate the flux is affected by gene manipulation (red), carbon supply (blue), both (green), and either (orange).

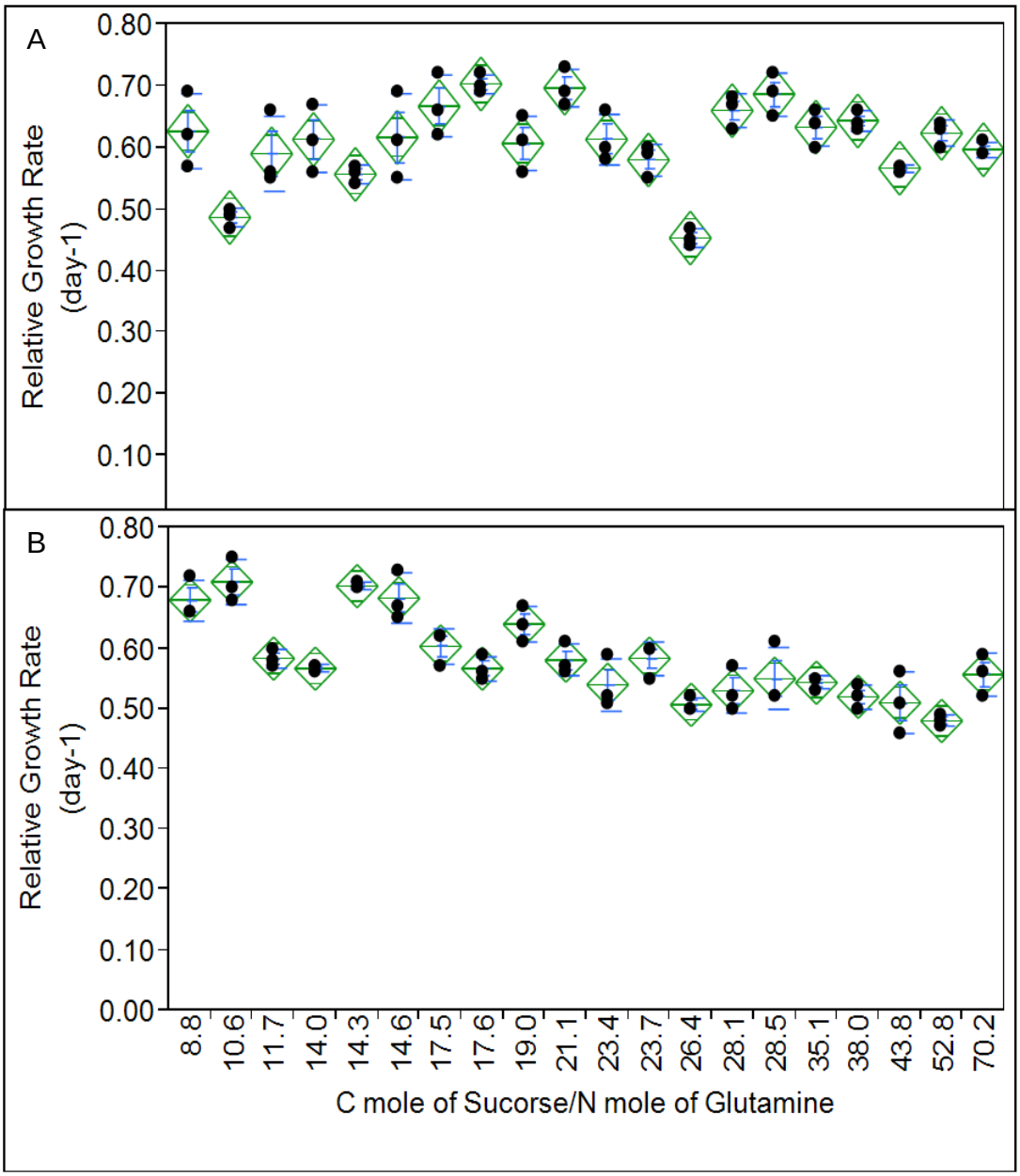

A
Relative Growth Rate (day-1)
0.80
0.70
0.60
0.50
0.40
0.30
0.20
0.10
B
Relative Growth Rate (day-1)
0.80
0.70
0.60
0.50
0.40
0.30
0.20
0.10
0.00
8.8
10.6
11.7
14.0
14.3
14.6
17.5
17.6
19.0
21.1
23.4
23.7
26.4
28.1
28.5
35.1
38.0
43.8
52.8
70.2
C mole of Sucorse/N mole of Glutamine

Figure 1

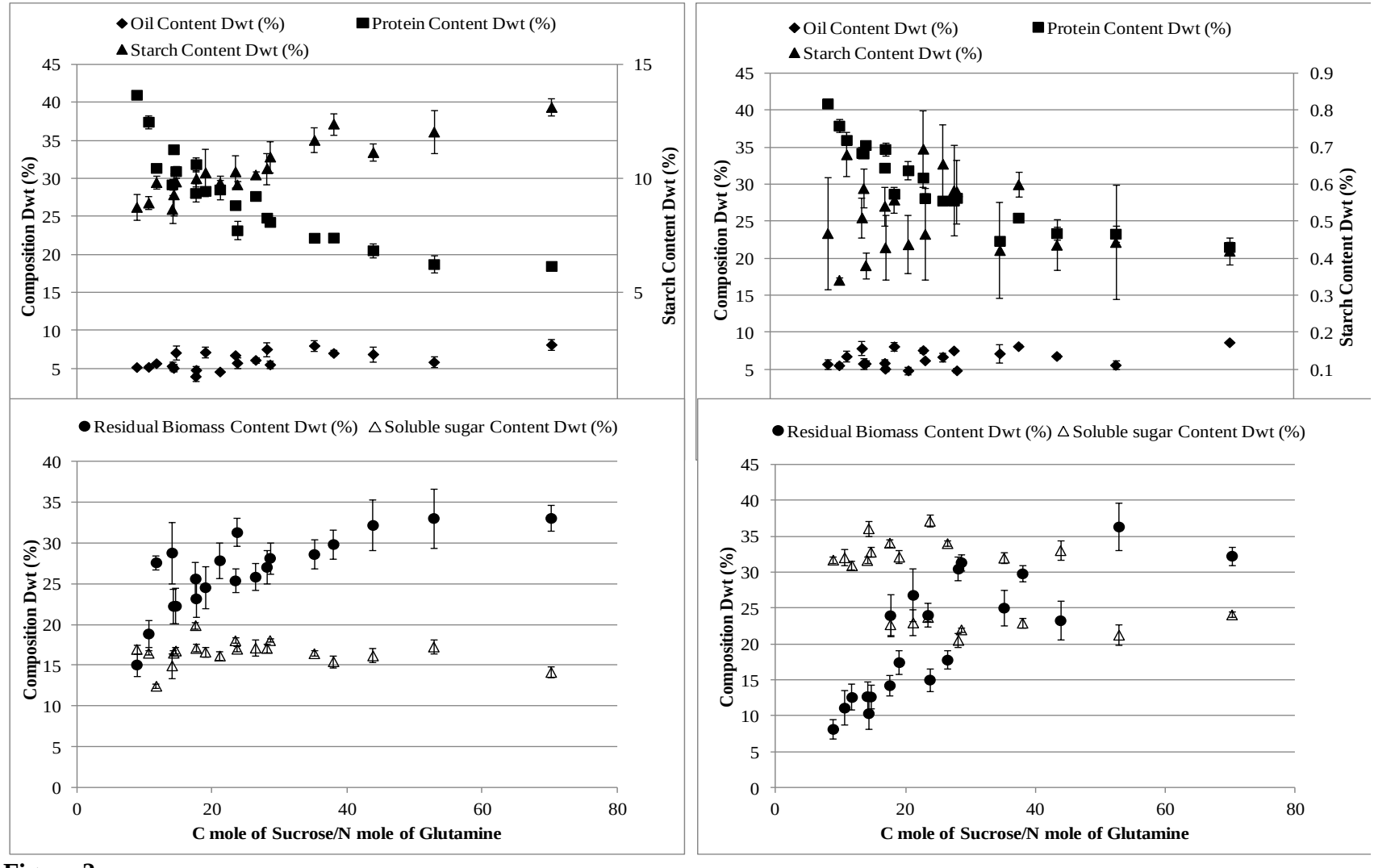

Oil Content Dwt (%)
Protein Content Dwt (%)
Starch Content Dwt (%)
Composition Dwt (%)
Starch Content Dwt (%)
Residual Biomass Content Dwt (%)
Soluble sugar Content Dwt (%)
C mole of Sucrose/N mole of Glutamine

Figure 2

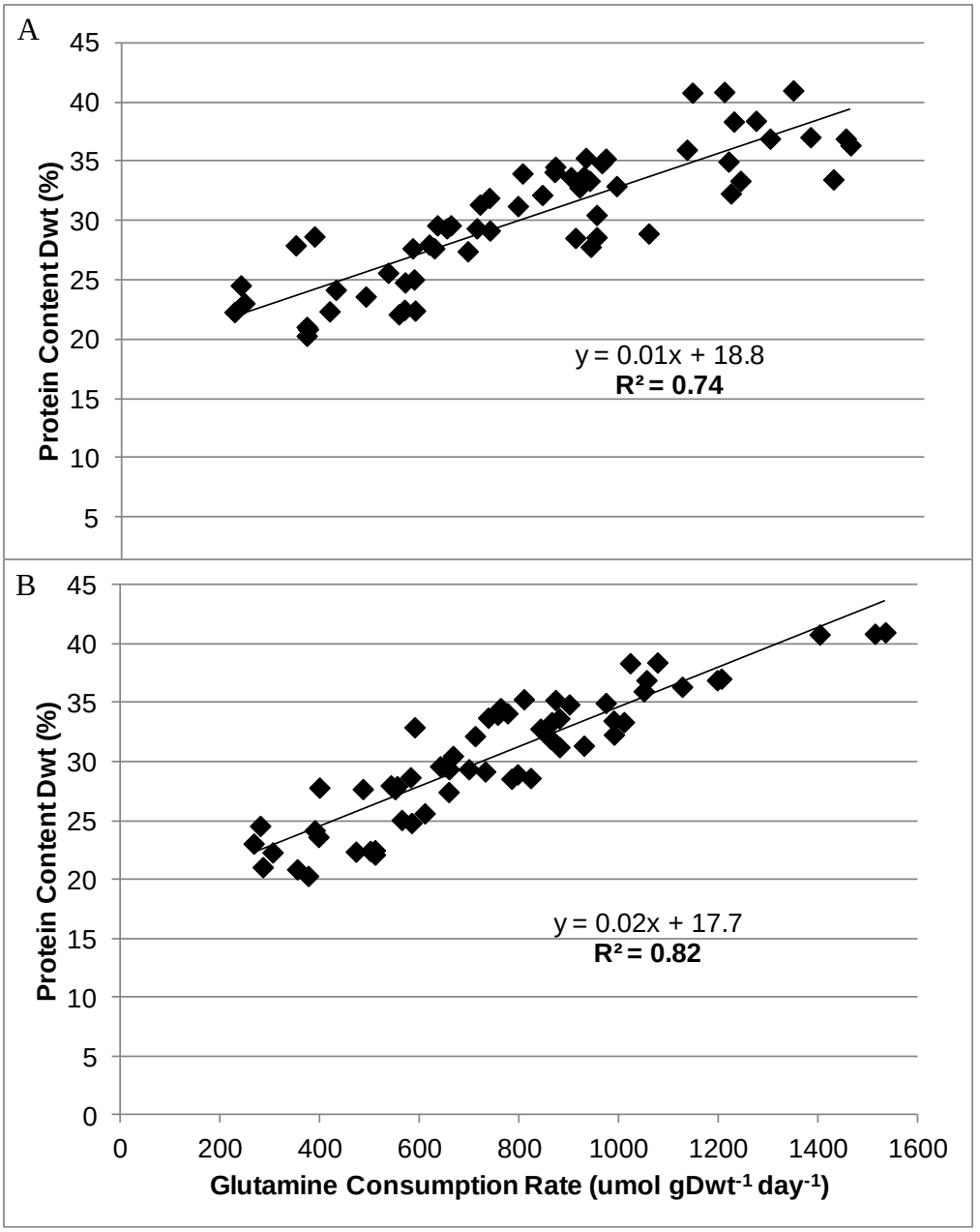

A
Protein Content Dwt (%)
y = 0.01x + 18.8
R² = 0.74
B
Protein Content Dwt (%)
y = 0.02x + 17.7
R² = 0.82
Glutamine Consumption Rate (umol gDwt-1 day-1)

Figure 3

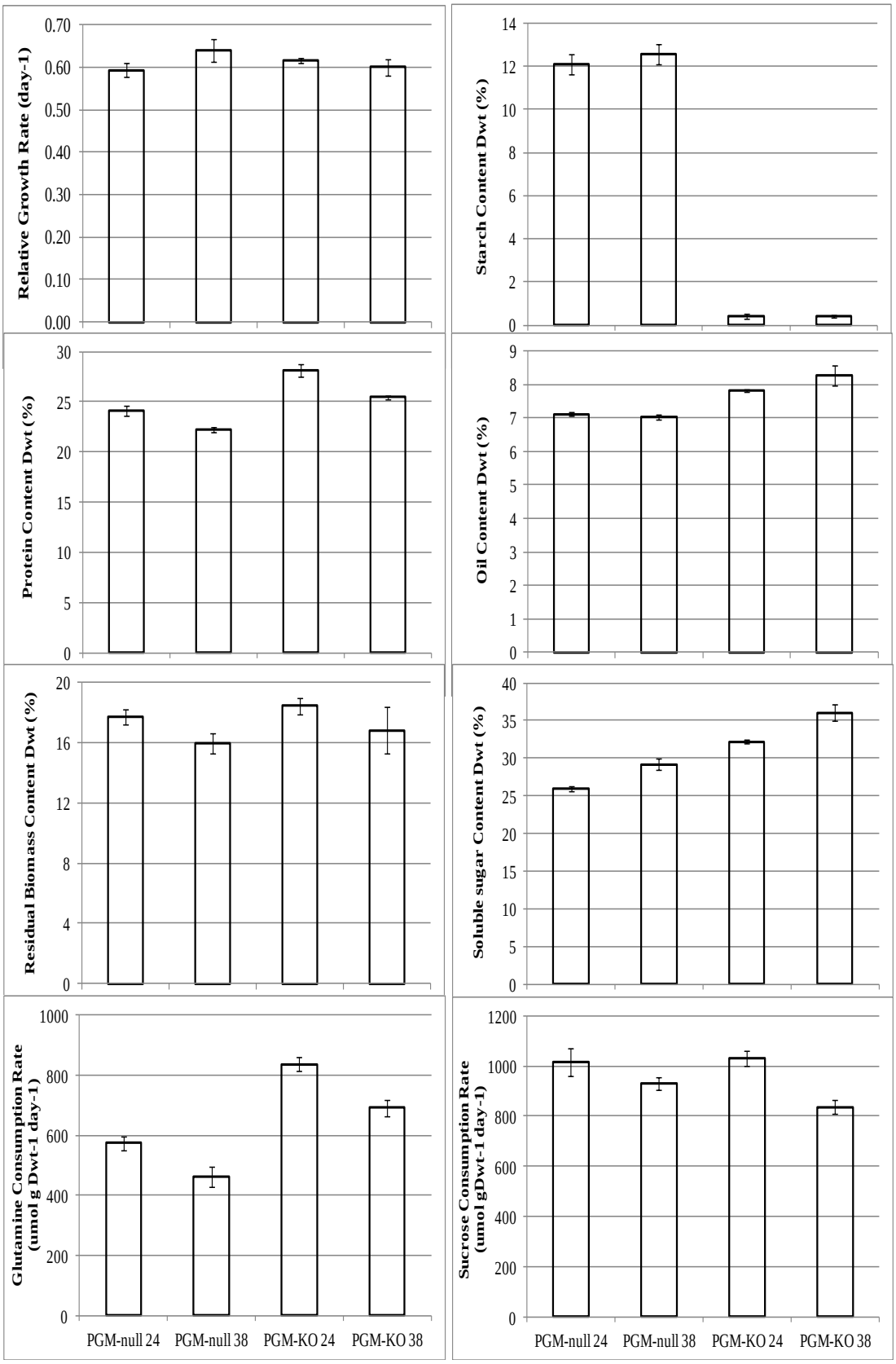

Relative Growth Rate (day-1)
0.70
0.60
0.50
0.40
0.30
0.20
0.10
0.00
Starch Content Dwt (%)
14
12
10
8
6
4
2
0
Protein Content Dwt (%)
30
25
20
15
10
5
0
Oil Content Dwt (%)
9
8
7
6
5
4
3
2
1
0
Residual Biomass Content Dwt (%)
20
16
12
8
4
0
Soluble sugar Content Dwt (%)
40
35
30
25
20
15
10
5
0
Glutamine Consumption Rate (umol g Dwt-1 day-1)
1000
800
600
400
200
0
Sucrose Consumption Rate (umol gDwt-1 day-1)
1200
1000
800
600
400
200
0
PGM-null 24
PGM-null 38
PGM-KO 24
PGM-KO 38

Figure 4

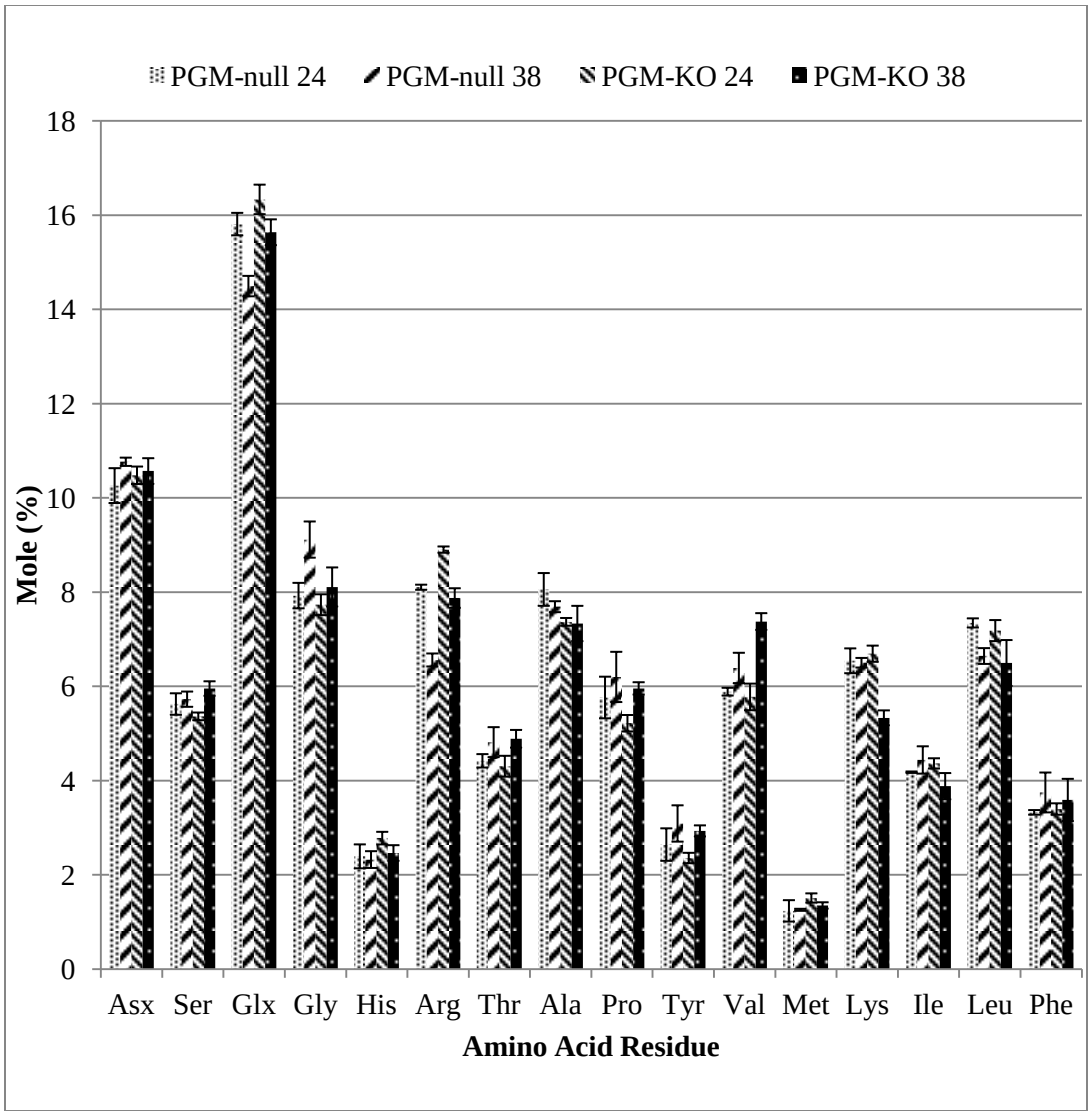

PGM-null 24
PGM-null 38
PGM-KO 24
PGM-KO 38
Mole (%)
0
2
4
6
8
10
12
14
16
18
Asx
Ser
Glx
Gly
His
Arg
Thr
Ala
Pro
Tyr
Val
Met
Lys
Ile
Leu
Phe
Amino Acid Residue

Figure 5

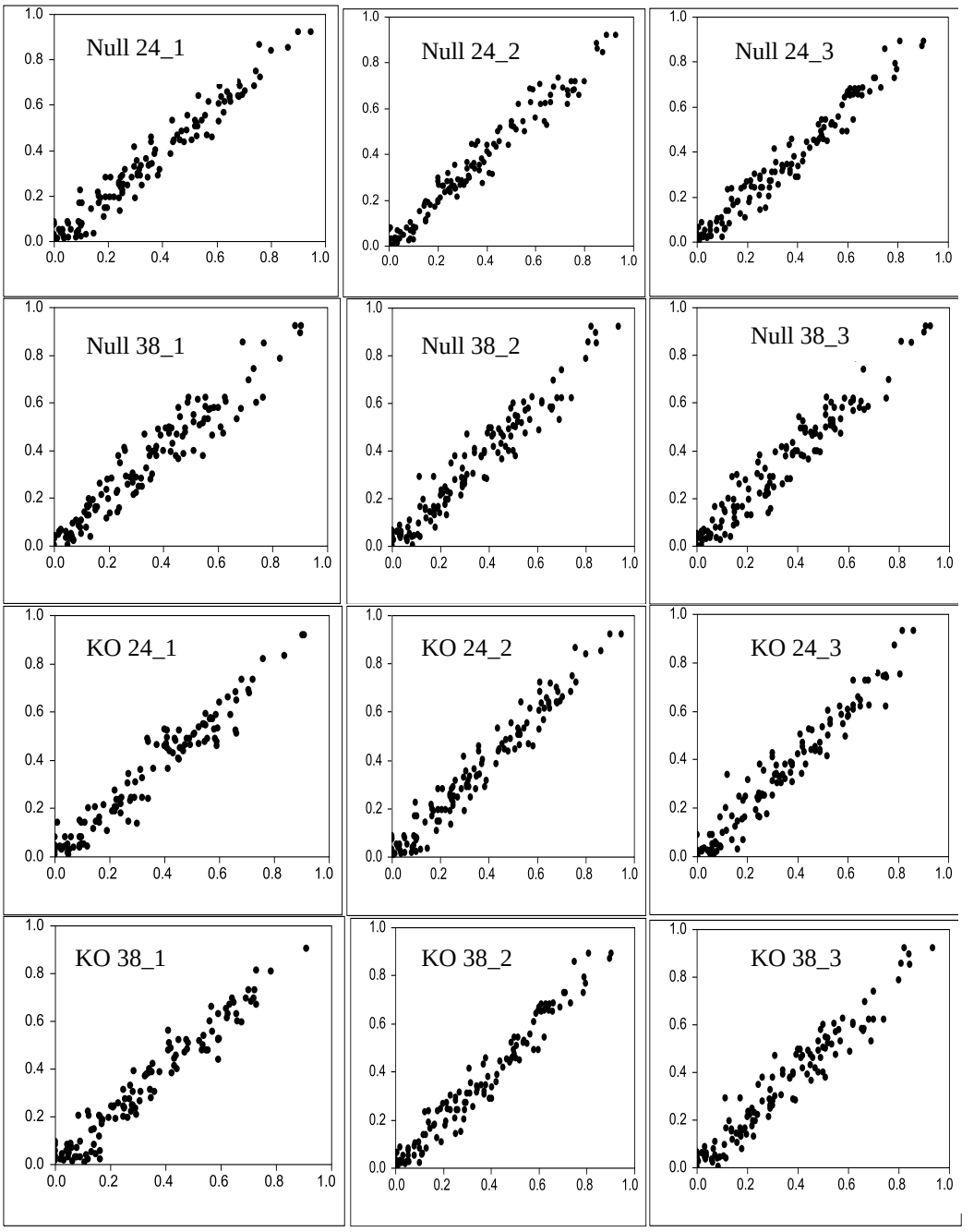

Null 24_1
Null 24_2
Null 24_3
Null 38_1
Null 38_2
Null 38_3
KO 24_1
KO 24_2
KO 24_3
KO 38_1
KO 38_2
KO 38_3
Simulated Value
Experimented Value

Figure 6

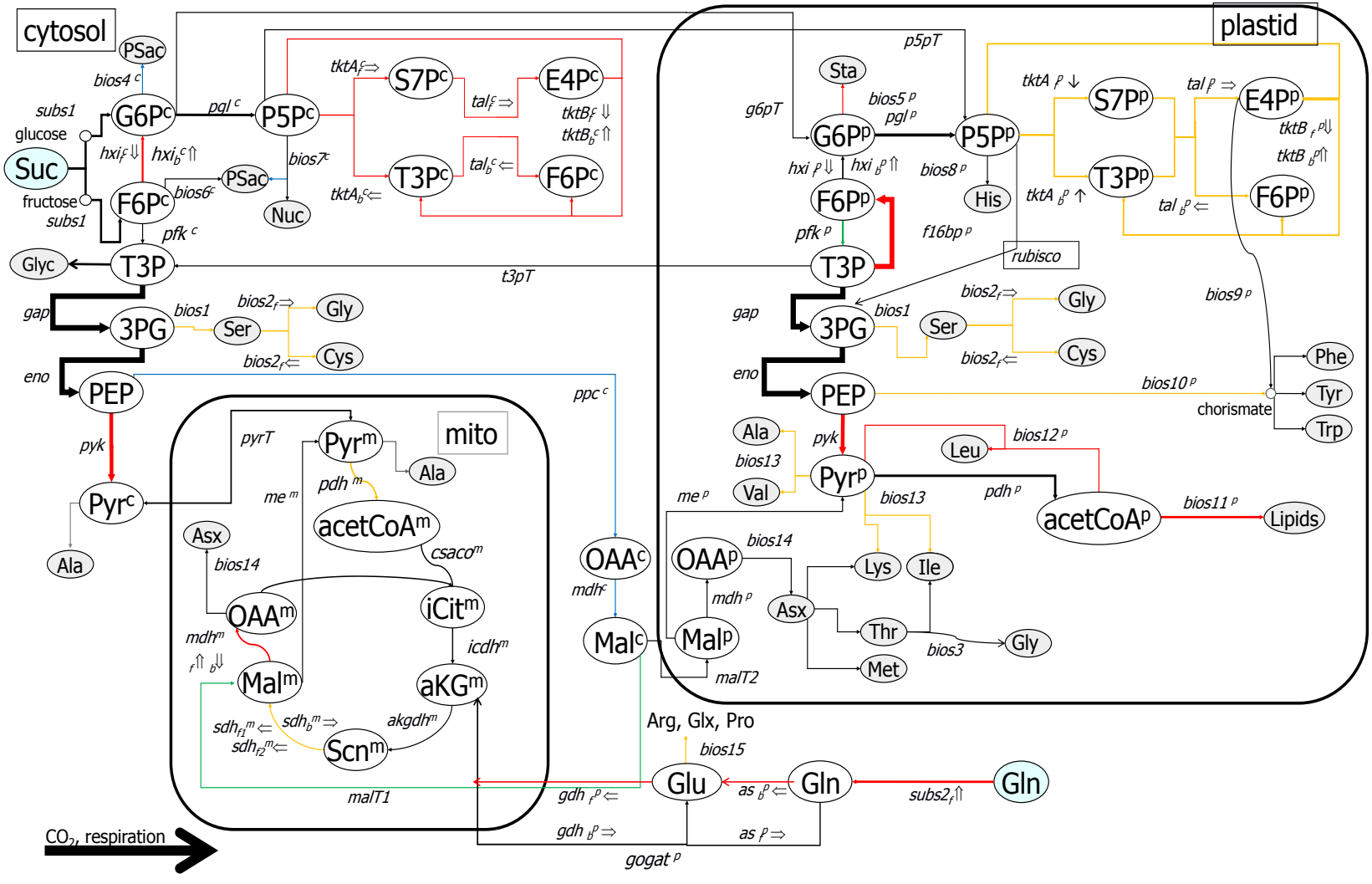

cytosol
plastid
mito
PSac
bios4 c
subs1
glucose
Suc
fructose
subs1
G6Pc
F6Pc
pgl c
P5Pc
hxi f c ⇓
hxi b c ⇑
bios6 c
bios7 c
PSac
Nuc
tktA f c ⇒
tktA b c ⇐
S7Pc
T3Pc
tal f c ⇒
tal b c ⇐
E4Pc
F6Pc
tktB f c ⇓
tktB b c ⇑
pfk c
Glyc
T3P
t3pT
gap
3PG
bios1
Ser
bios2 f ⇒
Gly
bios2 f ⇐
Cys
eno
PEP
pyk
Pyrc
Ala
ppc c
pyrT
Pyrm
Ala
me m
pdh m
acetCoAm
csaco m
Asx
bios14
OAAm
iCitm
mdh m
f ⇑ b ⇓
Malm
icdh m
akgdh m
aKGm
sdh f1 m ⇐
sdh f2 m ⇐
sdh b m ⇒
Scnm
malT1
OAAc
mdh c
Malc
p5pT
Sta
g6pT
G6Pp
P5Pp
pgl p
bios5 p
hxi f p ⇓
hxi b p ⇑
bios8 p
His
F6Pp
pfk p
f16bp p
rubisco
T3P
gap
3PG
bios1
Ser
bios2 f ⇒
Gly
bios2 f ⇐
Cys
eno
PEP
pyk
Pyrp
Ala
bios13
Val
me p
pdh p
acetCoAp
bios11 p
Lipids
Leu
bios12 p
bios13
OAAp
bios14
mdh p
Malp
malT2
Asx
Lys
Ile
Thr
Met
bios3
Gly
tktA f p ↓
S7Pp
T3Pp
tktA b p ↑
tal f p ⇒
E4Pp
tktB f p ⇓
tktB b p ⇑
tal b p ⇐
F6Pp
bios9 p
bios10 p
chorismate
Phe
Tyr
Trp
Arg, Glx, Pro
bios15
Glu
as b p ⇐
Gln
subs2 f ⇑
Gln
gdh f p ⇐
gdh b p ⇒
as f p ⇒
gogat p
CO2, respiration

Figure 7

Table 1: Media sucrose and glutamine concentrations and the mole ratios of carbon-nitrogen used in this study. Media conditions in boldface were chosen for 13C-metabolic flux analysis for soybean somatic embryos PGM-null and PGM-KO cultured in SHaM media.

C moles of Sucrose /N moles of Glutamine				
	Glutamine Conc. (mM) →			
Sucrose Conc. (mM) ↓	*20*	*37*	*50*	*60*
88	26.4	14.3	10.6	8.8
117	35.1	19.0	14.0	11.7
146	43.8	**23.7**	17.5	14.6
176	52.8	28.5	21.1	17.6
234	70.2	**38.0**	28.1	23.4

Table 2: Average values of relative fluxes (normalized to 100 mole of sucrose uptake) for the four PGM cultures analyzed in this study. For abbreviations of flux names, see Figure 7. Values show best fit of flux estimation with standard deviation as determined based on Monte Carlo simulation of 500 times (see "Materials & Methods"). Each net flux with standard deviation represents three separate biological replicates; in total, 12 cultures were examined and 12 flux maps were developed. Boldface values indicate that the difference of fluxes between the PGM-null 24, PGM-null 38, PGM-KO 24, and PGM-KO 38 are significant (P = 90%; Student t' test mean comparison with n 3 for all of the PGM cultures).

Reaction	Reaction *Name*	PGM-null 24 Average		PGM-null 38 Avearge		PGM-KO 24 Average		PGM-KO 38 Average	
		Net Flux	S.D.	Net Flux	S.D.	Net Flux	S.D.	Net Flux	S.D.
Glycolysis and oxPPP									
$G6P^c \leftrightarrow F6P^c$	hxi_f^c	**-53**	**17**	**-34**	**8**	**-103**	**11**	**-112**	**9**
reversibility %		0.33	0.24	0.77	0.14	0.42	0.14	0.47	0.18
$G6P^p \leftrightarrow F6P^p$	hxi_f^p	-74	24	-32	8	-44	25	-38	17
reversibility %		0.23	0.24	0.21	0.15	0.69	0.21	0.31	0.23
$G6P^c \rightarrow P5P^c + CO2$	pgl^c	112	28	45	9	137	26	107	42
$G6P^p \rightarrow P5P^p + CO2$	pgl^p	92	33	48	16	69	26	115	43
$P5P^c + P5P^c \leftrightarrow S7P^c + T3P^c$	$tktA_f^c$	**6**	**2**	**5**	**1**	**29**	**6**	**21**	**6**
reversibility %		0.72	0.04	0.8	0.02	0.76	0.2	0.71	0.17
$S7P^c + T3P^c \leftrightarrow F6P^c + E4P^c$	tal_f^c	**6**	**2**	**5**	**1**	**29**	**6**	**21**	**6**
reversibility %		0.82	0.2	0.96	0.01	0.72	0.16	0.89	0.09
$P5P^c + E4P^c \leftrightarrow F6P^c + T3P^c$	$tktB_f^c$	**6**	**2**	**5**	**1**	**29**	**6**	**21**	**6**
reversibility %		0.85	0.11	0.69	0.23	0.69	0.1	0.64	0.22
$P5P^p + P5P^p \leftrightarrow S7P^p + T3P^p$	$tktA_f^p$	**55**	**7**	**19**	**4**	**20**	**7**	**30**	**8**
reversibility %		0.84	0.16	0.95	0.01	0.77	0.23	0.48	0.36
$S7P^p + T3P^p \leftrightarrow F6P^p + E4P^p$	tal_f^p	**55**	**7**	**19**	**4**	**20**	**7**	**30**	**8**
reversibility %		0.86	0.13	0.98	0.01	0.62	0.18	0.7	0.29
$P5P^p + E4P^p \leftrightarrow F6P^p + T3P^p$	$tktB_f^p$	**53**	**7**	**17**	**4**	**25**	**7**	**27**	**8**
reversibility %		0.98	0.01	0.86	0.1	0.71	0.2	0.98	0.02

Table 2: (continued.)

Reaction								Reaction Name	PGM-null 24 Average		PGM-null 38 Averge		PGM-KO 24 Average		PGM-KO 38 Average	
									Net Flux	S.D.	Net Flux	S.D.	Net Flux	S.D.	Net Flux	S.D.
Glycolysis and oxPPP																
		$F6P^c$	→	$T3P^c$	+	$T3P^c$		*pfk*^{c}	39	14	4	3	12	10	13	10
		$F6P^p$	→	$T3P^p$	+	$T3P^p$		*pfk*^{p}	**33**	**13**	**95**	**2**	**83**	**12**	**50**	**12**
$T3P^p$	+	$T3P^p$	→	$F6P^p$				*f16bp*^{p}	**107**	**7**	**127**	**10**	**73**	**7**	**57**	**10**
		T3P	→	3PG				*gap*	199	17	213	13	226	18	197	20
		3PG	→	PEP				*eno*	194	17	208	13	221	18	185	13
		PEP	→	Pyr				*pyk*	**108**	**19**	**63**	**12**	**156**	**18**	**116**	**17**
		Pyr^p	→	ACA^p	+	CO_2		*pdh*^{p}	77	12	85	1	99	14	76	17
		Pyr^m	→	ACA^m	+	CO_2		*pdh*^{m}	**44**	**1**	**47**	**1**	**45**	**2**	**53**	**1**
		Pyr^c	→	Pyr^m				*pyrT*	53	7	59	2	53	6	59	5
		$G6P^c$	↔	$G6P^p$				$g6pT_f$	32	17	33	10	40	21	75	43
		reversibility %							0.34	0.22	0.2	0.17	0.44	0.21	0.41	0.3
		$P5P^c$	↔	$P5P^p$				$p5pT_f$	72	26	8	8	20	26	55	23
		reversibility %							0.65	0.22	0.64	0.22	0.57	0.2	0.62	0.23
$P5P^p$	+	CO_2	→	3PG	+	3PG		*rubisco*	1	1	1	1	6	5	12	7
		$T3P^c$	↔	$T3P^p$				$t3pT_f$	26	16	12	5	20	12	34	26
		reversibility %							0.29	0.23	0.2	0.17	0.66	0.14	0.83	0.24
TCA cycle																
ACA^m	+	OAA^m	→	$ICit^m$				*csaco*^{m}	38	12	42	3	50	14	38	17
		$ICit^m$	→	aKG^m	+	CO_2		*icdh*^{m}	38	12	42	3	50	14	38	17
		aKG^m	→	Scn^m	+	CO_2		*akgdh*^{m}	32	13	35	3	42	14	34	18
		Scn^m	→	Mal^m				$sdh_{f1}{}^m$	32	13	35	3	42	14	34	18
		Scn^m	↔	Mal^m				$sdh_{f2}{}^m$	**27**	**6**	**32**	**3**	**31**	**5**	**11**	**2**

Table 2: (continued.)

Reaction	Reaction Name	PGM-null 24 Average		PGM-null 38 Average		PGM-KO 24 Average		PGM-KO 38 Average	
		Net Flux	S.D.	Net Flux	S.D.	Net Flux	S.D.	Net Flux	S.D.
reversibility %		0.71	0.06	0.98	0.01	0.74	0.19	0.78	0.3
$Mal^m \leftrightarrow OAA^m$	mdh_f^m	41	12	45	3	52	14	41	17
reversibility %		0.49	0.37	0.48	0.36	0.55	0.35	0.51	0.33
Anaplerotic reactions									
$PEP^c + CO_2 \rightarrow OAA^c$	ppc^c	**29**	**4**	**48**	**1**	**22**	**4**	**25**	**5**
$Mal^m \rightarrow Pyr^m + CO_2$	me^m	8	6	12	1	8	4	6	3
$Mal^p \rightarrow Pyr^p + CO_2$	me^p	8	4	12	1	9	2	10	2
Substrate entry									
$Suc^{ext} \rightarrow G6P^c + F6P^c$	*subs1*	100	0	100	0	100	0	100	0
$Gln^{ext} \leftrightarrow Gln^p$	$subs2_f$	**46**	**2**	**47**	**1**	**66**	**3**	**61**	**2**
reversibility %		0.9	0.01	0.76	0.01	0.74	0.19	0.93	0.05
Glutamate assimilation									
$Gln^p \rightarrow Glu^p$	as_f	**58**	**9**	**70**	**7**	**41**	**3**	**26**	**5**
$Glu^p \rightarrow aKG^p$	gdh_f^p	**59**	**9**	**72**	**7**	**85**	**4**	**37**	**5**
$aKG^p + Gln^p \rightarrow Glu^p + Glu^p$	$gogat^p$	29	9	36	7	31	6	21	7
Malate shuttle									
$Mal^m \leftrightarrow Mal^c$	$malT1_f$	**-11**	**1**	**-23**	**2**	**1**	**1**	**3**	**2**
reversibility %		0.91	0.09	0.9	0	0.97	0.01	0.95	0.02
$Mal^c \rightarrow Mal^p$	*malT2*	22	4	26	2	21	3	25	3

Table 2: (continued.)

Mal^c	→	OAA^c		*mdh^c*	**29**	**4**	**48**	**1**	**22**	**4**	**25**	**5**
Mal^p	→	OAA^p		*mdh^p*	13	1	14	1	14	1	15	1
Biosynthesis of Ser and Gly												
3PG	→	Ser		*bios1*	**4**	**0**	**5**	**0**	**5**	**0**	**6**	**0**
Ser	↔	Gly	+ C1	*$bios2_f$*	**4**	**0**	**5**	**0**	**5**	**0**	**6**	**0**
reversibility %					0.43	0.04	0.38	0.01	0.39	0.04	0.46	0.03
Thr	→	Gly		*bios3*	3	1	3	0	2	1	2	1
Fluxes towards biosynthesis (other than those related to Ser and Gly) and effluxes into medium												
CO_2	→			*resp*	**380**	**26**	**354**	**6**	**457**	**22**	**380**	**26**
$G6P^c$	→	biomass		*bios4*	**27**	**1**	**33**	**1**	**27**	**1**	**34**	**1**
$G6P^p$	→	biomass		*bios5*	**15**	**1**	**16**	**0**	**0**	**0**	**0**	**0**
$F6P^c$	→	biomass		*bios6*	3	0	3	0	3	0	3	0
$P5P^c$	→	biomass		*bios7*	**21**	**1**	**23**	**0**	**28**	**1**	**33**	**1**
$P5P^p$	→	biomass		*bios8*	1	0	1	0	1	0	1	0
$E4P^p$	→	biomass		*bios9*	2	0	2	0	2	0	2	0
PEP^p	→	biomass		*bios10*	**4**	**0**	**5**	**0**	**4**	**0**	**8**	**0**
ACA^p	→	biomass		*bios11*	**86**	**1**	**87**	**1**	**89**	**1**	**102**	**1**
ACA^p	→	biomass		*bios12*	**2**	**0**	**2**	**0**	**3**	**0**	**3**	**0**
Glycerol	→	biomass		*glyc*	2	1	3	1	2	1	2	1
Pyr	→	biomass		*bios13*	**15**	**1**	**14**	**2**	**16**	**1**	**25**	**3**
OAA	→	biomass		*bios14*	13	1	14	1	12	1	12	1
Glu	→	biomass		*bios15*	**10**	**0**	**10**	**0**	**12**	**1**	**18**	**1**
				Chisquare	100	22	112	33	79	15	76	14

Supplemental Data

Figure S1: Flux toward carbohydrate biosynthesis

Figure S2: Flux toward plastic and cytosolic pentose phosphate pathway

Figure S3: Flux toward glycolysis and lipid biosynthesis

Figure S4: Glutamine uptake flux and glutamate assimilation fluxes

Figure S5: Flux toward protein biosynthesis

Figure S6: Fluxes from TCA cycle, anaplerotic reaction, and malate shuttles

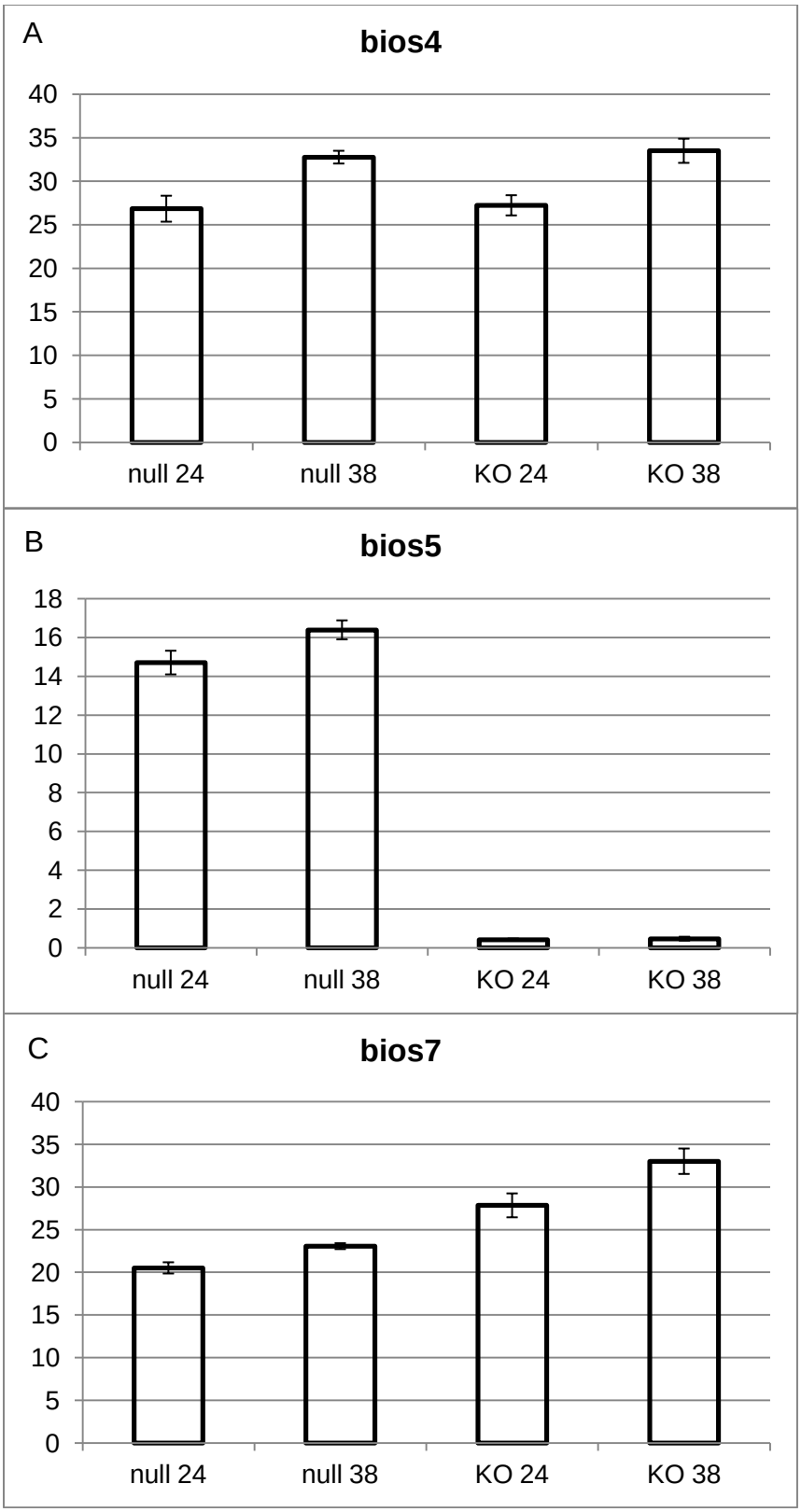

A
bios4
40
35
30
25
20
15
10
5
0
null 24
null 38
KO 24
KO 38
B
bios5
18
16
14
12
10
8
6
4
2
0
null 24
null 38
KO 24
KO 38
C
bios7
40
35
30
25
20
15
10
5
0
null 24
null 38
KO 24
KO 38
Relative flux (normalized to 100 mole of sucrose uptake)

Figure S1

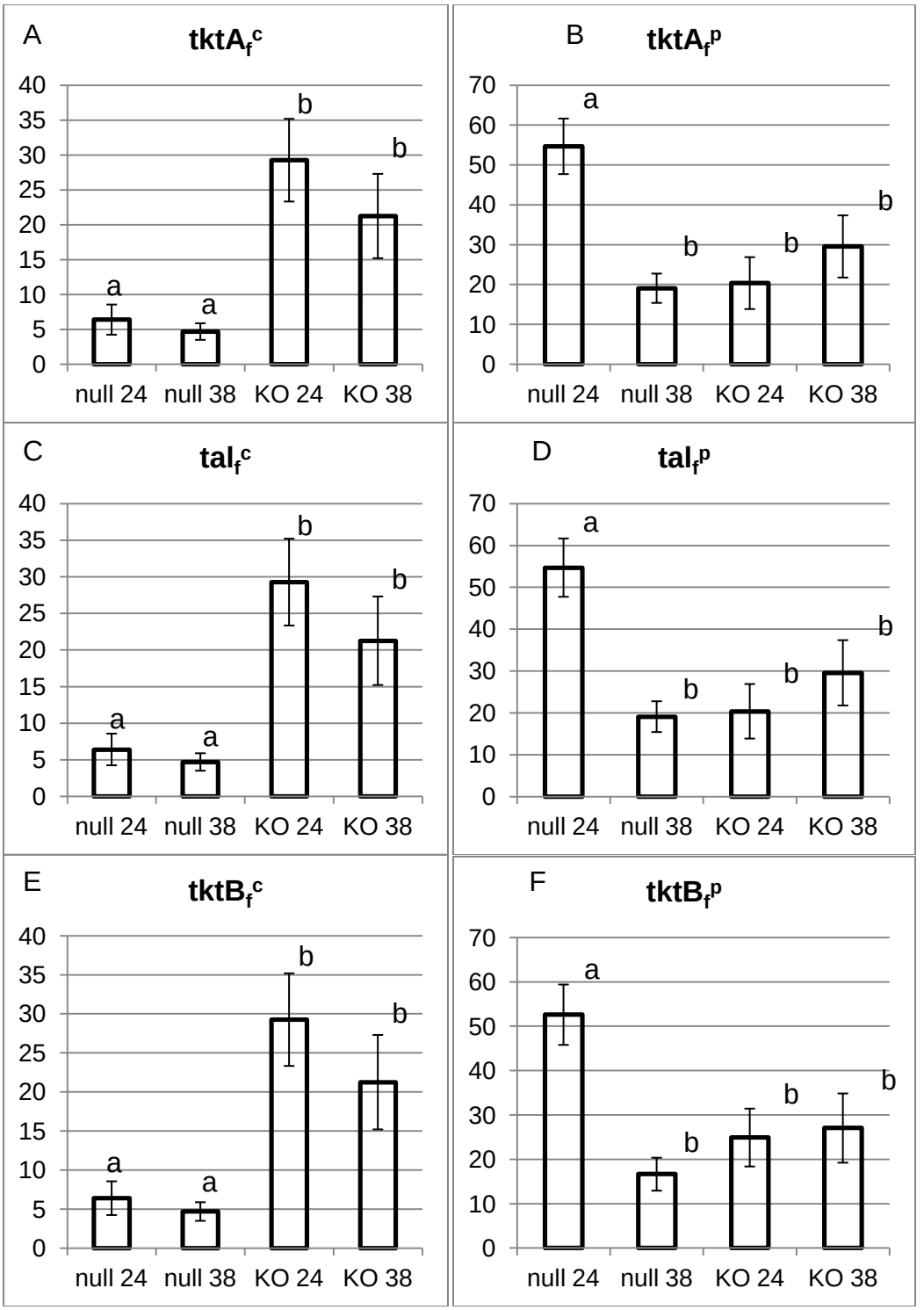

A
tktA$_f^c$
B
tktA$_f^p$
C
tal$_f^c$
D
tal$_f^p$
E
tktB$_f^c$
F
tktB$_f^p$
null 24
null 38
KO 24
KO 38
Relative flux (normalized to 100 mole of sucrose uptake)

Figure S2

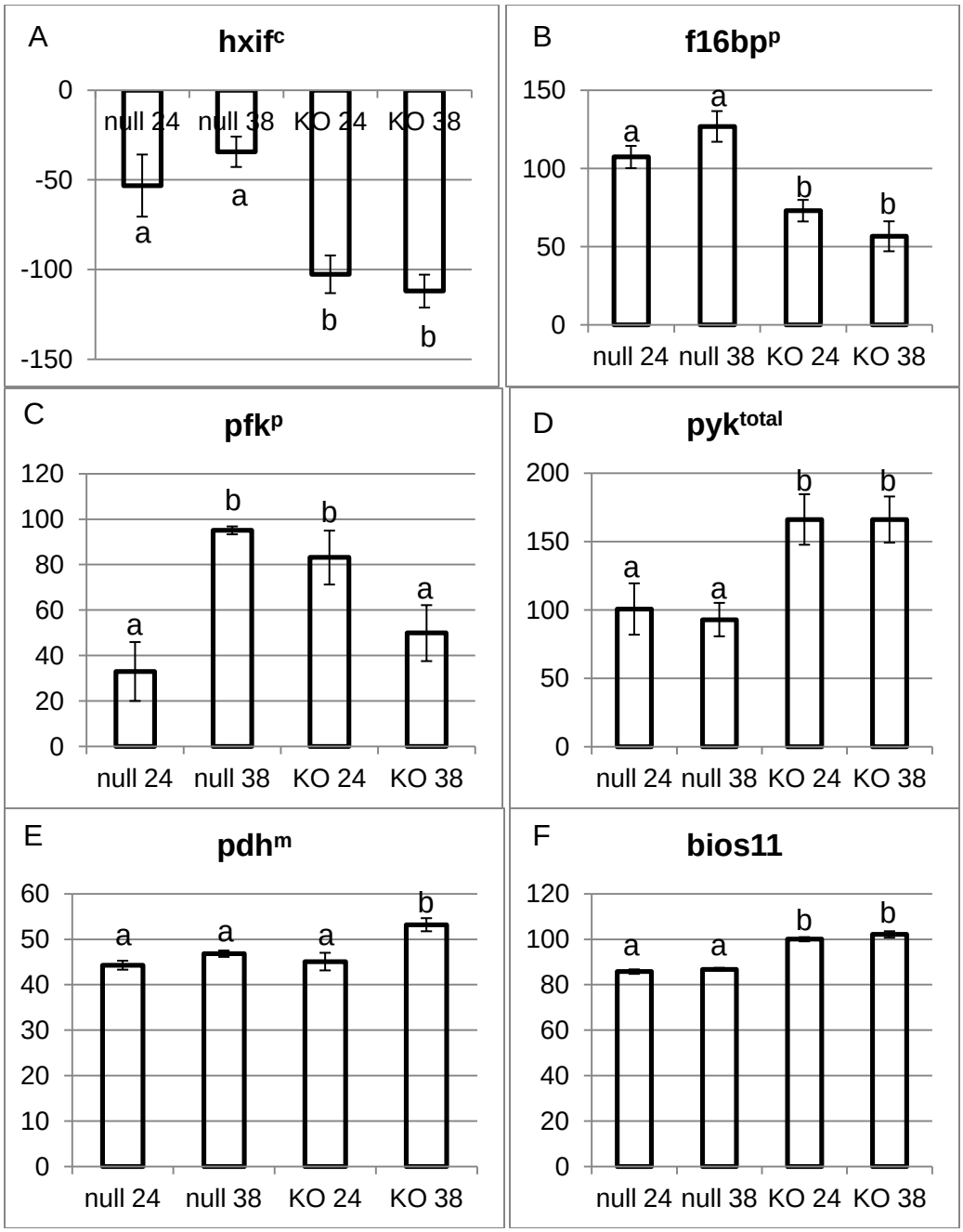

A
hxifc
0
-50
-100
-150
null 24
null 38
KO 24
KO 38
a
a
b
b
B
f16bpp
150
100
50
0
a
a
b
b
C
pfkp
120
100
80
60
40
20
0
b
b
a
a
D
pyktotal
200
150
100
50
0
a
a
b
b
E
pdhm
60
50
40
30
20
10
0
a
a
a
b
F
bios11
120
100
80
60
40
20
0
a
a
b
b
Relative flux (normalized to 100 mole of sucrose uptake)

Figure S3

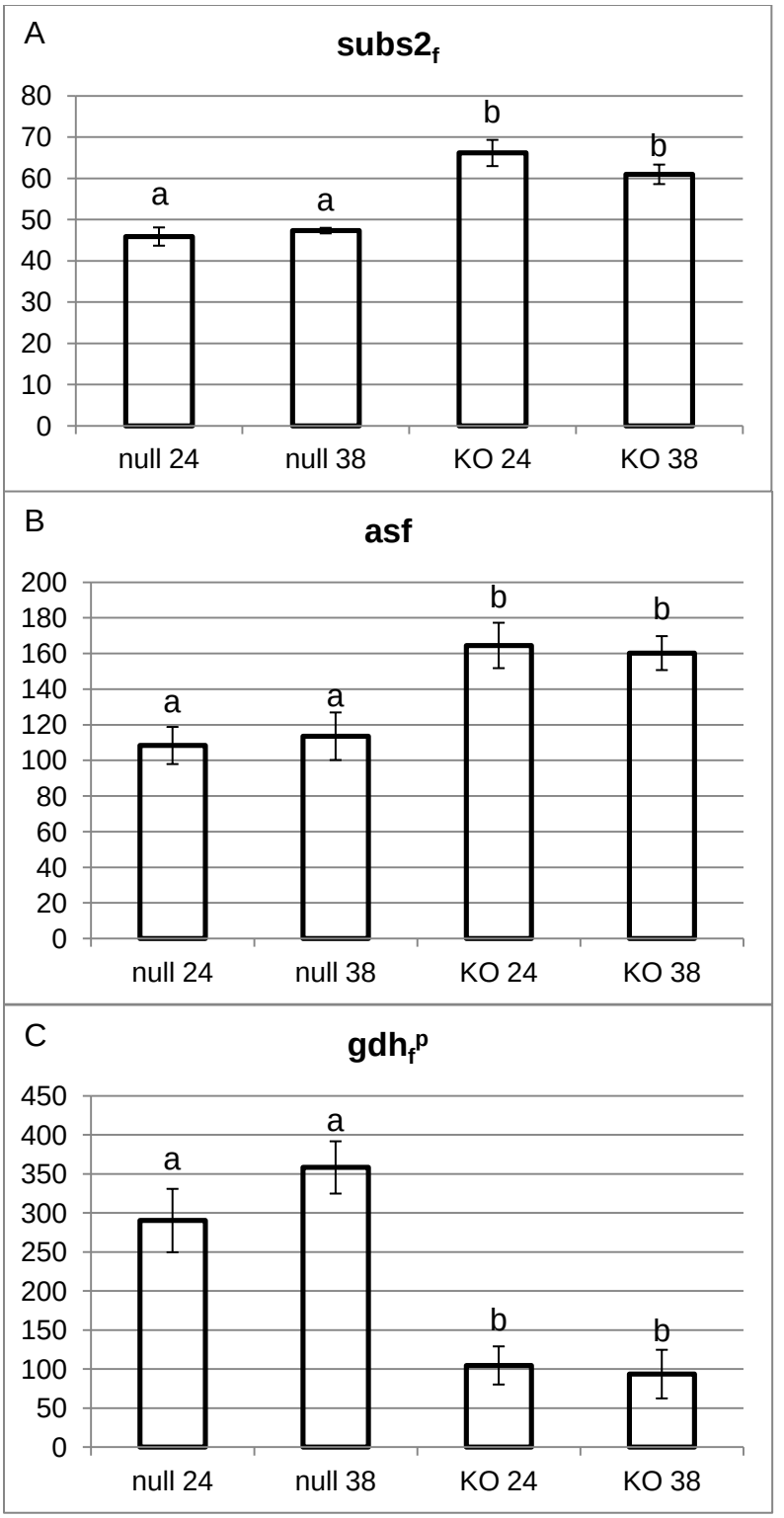

A
subs2f
80
70
60
50
40
30
20
10
0
a
a
b
b
null 24
null 38
KO 24
KO 38
B
asf
200
180
160
140
120
100
80
60
40
20
0
a
a
b
b
null 24
null 38
KO 24
KO 38
C
gdhfp
450
400
350
300
250
200
150
100
50
0
a
a
b
b
null 24
null 38
KO 24
KO 38
Relative flux (normalized to 100 mole of sucrose uptake)

Figure S4

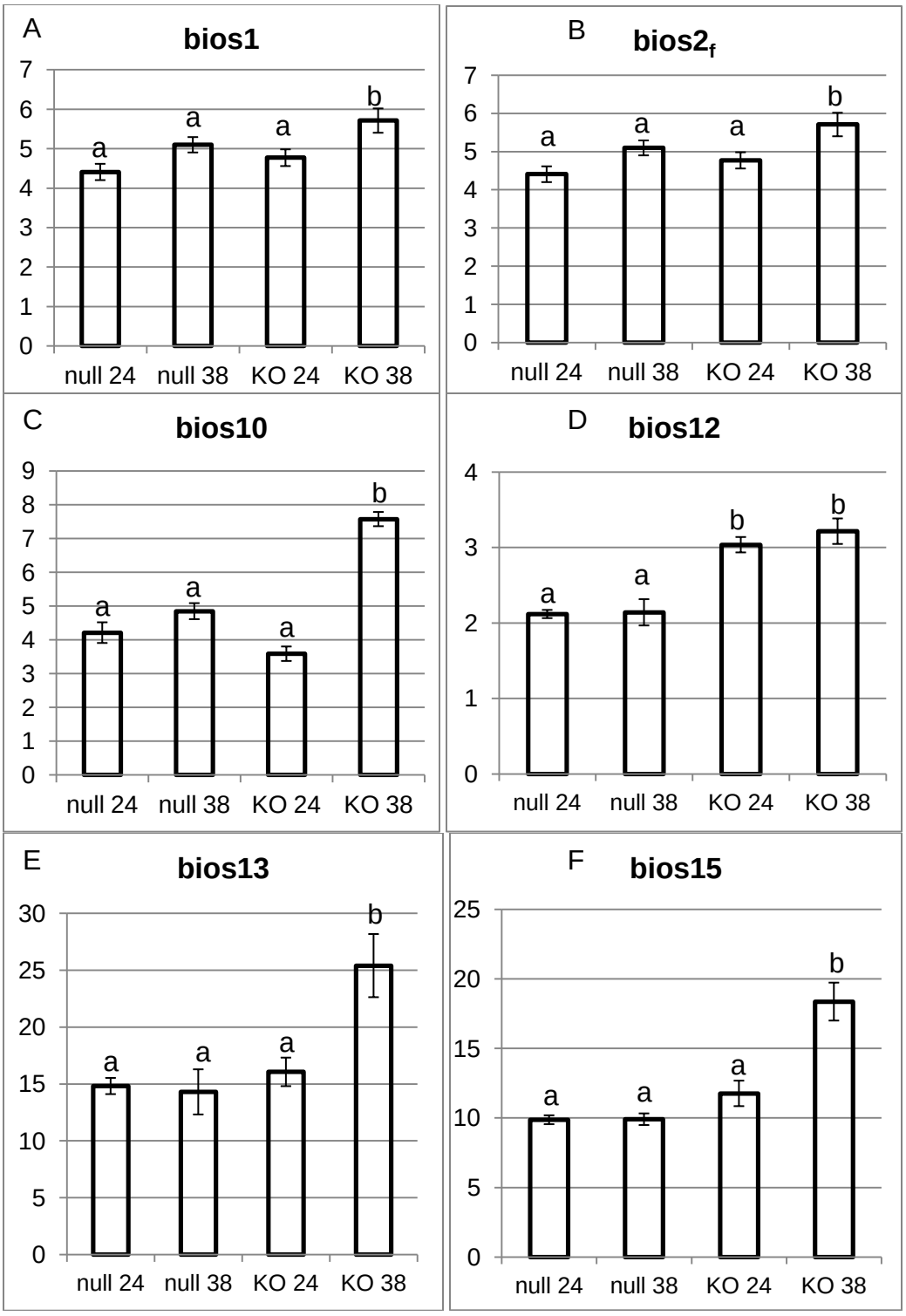

A
bios1
B
bios2f
C
bios10
D
bios12
E
bios13
F
bios15
null 24
null 38
KO 24
KO 38
Relative flux (normalized to 100 mole of sucrose uptake)

Figure S5

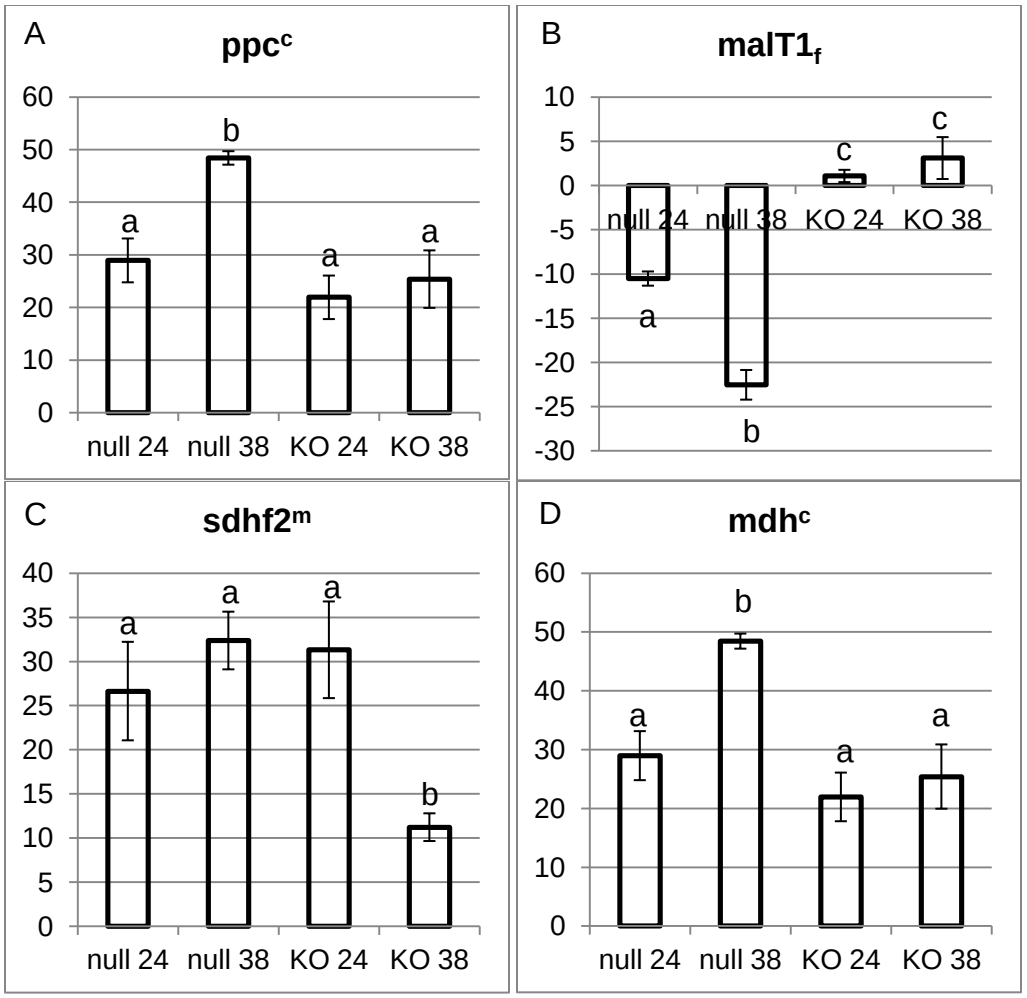

A
ppc^c
60
50
40
30
20
10
0
a
b
a
a
null 24
null 38
KO 24
KO 38
B
malT1_f
10
5
0
-5
-10
-15
-20
-25
-30
a
b
c
c
null 24
null 38
KO 24
KO 38
C
sdhf2^m
40
35
30
25
20
15
10
5
0
a
a
a
b
null 24
null 38
KO 24
KO 38
D
mdh^c
60
50
40
30
20
10
0
a
b
a
a
null 24
null 38
KO 24
KO 38
Relative flux (normalized to 100 mole of sucrose uptake)

Figure S6

Table S1: Average values of absolute fluxes (umol metabolite gDwt-1 day-1) for the four PGM cultures analyzed in this study. For abbreviations of flux names, see Figure 7. Values show best fit of flux estimation with standard deviation as determined based on Monte Carlo simulation of 500 times (see "Materials & Methods"). Each net flux with standard deviation represents three separate biological replicates; in total, 12 cultures were examined and 12 flux maps were developed. Boldface values indicate that the difference of fluxes between the PGM-null 24, PGM-null 38, PGM-KO 24, and PGM-KO 38 are significant (P = 90%; Student t' test mean comparison with n 3 for all of the PGM cultures).

Reaction	Reaction *Name*	PGM-null 24 Average		PGM-null 38 Avearge		PGM-KO 24 Average		PGM-KO 38 Average	
		Net Flux	S.D.	Net Flux	S.D.	Net Flux	S.D.	Net Flux	S.D.
Glycolysis and oxPPP									
$G6P^c \leftrightarrow F6P^c$	hxi_f^c	**-539**	**175**	**-312**	**77**	**-1057**	**108**	**-935**	**77**
reversibility %		0.33	0.24	0.77	0.14	0.42	0.14	0.47	0.18
$G6P^p \leftrightarrow F6P^p$	hxi_f^p	-752	248	-288	73	-454	260	-316	146
reversibility %		0.23	0.24	0.21	0.15	0.69	0.21	0.31	0.23
$G6P^c \rightarrow P5P^c + CO2$	pgl^c	1131	279	412	83	1409	270	895	352
$G6P^p \rightarrow P5P^p + CO2$	pgl^p	928	335	438	150	712	268	959	359
$P5P^c + P5P^c \leftrightarrow S7P^c + T3P^c$	$tktA_f^c$	**65**	**22**	**43**	**11**	**301**	**61**	**177**	**51**
reversibility %		0.72	0.04	0.8	0.02	0.76	0.2	0.71	0.17
$S7P^c + T3P^c \leftrightarrow F6P^c + E4P^c$	tal_f^c	**65**	**22**	**43**	**11**	**301**	**61**	**177**	**51**
reversibility %		0.82	0.2	0.96	0.01	0.72	0.16	0.89	0.09
$P5P^c + E4P^c \leftrightarrow F6P^c + T3P^c$	$tktB_f^c$	**65**	**22**	**43**	**11**	**301**	**61**	**177**	**51**
reversibility %		0.85	0.11	0.69	0.23	0.69	0.1	0.64	0.22
$P5P^p + P5P^p \leftrightarrow S7P^p + T3P^p$	$tktA_f^p$	**554**	**71**	**173**	**33**	**210**	**67**	**247**	**65**
reversibility %		0.84	0.16	0.95	0.01	0.77	0.23	0.48	0.36
$S7P^p + T3P^p \leftrightarrow F6P^p + E4P^p$	tal_f^p	**554**	**71**	**173**	**33**	**210**	**67**	**247**	**65**
reversibility %		0.86	0.13	0.98	0.01	0.62	0.18	0.7	0.29

Table S1: (continued.)

$P5P^p$	+	$E4P^p$	↔	$F6P^p$	+	$T3P^p$	$tktB_f^p$	**533**	**69**	**151**	**34**	**257**	**67**	**226**	**65**
		reversibility %						0.98	0.01	0.86	0.1	0.71	0.2	0.98	0.02

Reaction							Reaction	PGM-null 24		PGM-null 38		PGM-KO 24		PGM-KO 38	
							Name	Average		Avearge		Average		Average	
								Net Flux	S.D.	Net Flux	S.D.	Net Flux	S.D.	Net Flux	S.D.
Glycolysis and oxPPP															
		$F6P^c$	→	$T3P^c$	+	$T3P^c$	pfk^c	400	139	37	27	120	98	111	84
		$F6P^p$	→	$T3P^p$	+	$T3P^p$	pfk^p	**334**	**131**	**864**	**16**	**856**	**122**	**417**	**103**
$T3P^p$	+	$T3P^p$	→	$F6P^p$			$f16bp^p$	**1087**	**72**	**1152**	**89**	**752**	**71**	**473**	**80**
		T3P	→	3PG			*gap*	**2012**	**177**	**1934**	**119**	**2328**	**189**	**1646**	**165**
		3PG	→	PEP			*eno*	**1967**	**175**	**1887**	**117**	**2278**	**187**	**1543**	**112**
		PEP	→	Pyr			*pyk*	**1089**	**190**	**570**	**111**	**1607**	**190**	**968**	**141**
		Pyr^p	→	ACA^p	+	CO_2	pdh^p	775	125	770	13	1024	141	637	143
		Pyr^m	→	ACA^m	+	CO_2	pdh^m	449	10	425	6	464	20	444	12
		Pyr^c	→	Pyr^m			*pyrT*	535	68	534	17	543	59	496	40
		$G6P^c$	↔	$G6P^p$			$g6pT_f$	325	168	299	90	409	216	629	359
		reversibility %						0.34	0.22	0.2	0.17	0.44	0.21	0.41	0.3
		$P5P^c$	↔	$P5P^p$			$p5pT_f$	728	266	75	76	211	265	462	188
		reversibility %						0.65	0.22	0.64	0.22	0.57	0.2	0.62	0.23
$P5P^p$	+	CO_2	→	3PG	+	3PG	*rubisco*	7	10	7	8	62	46	97	59
		$T3P^c$	↔	$T3P^p$			$t3pT_f$	264	158	106	45	210	122	282	217
		reversibility %						0.29	0.23	0.2	0.17	0.66	0.14	0.83	0.24
TCA cycle															
ACA^m	+	OAA^m	→	$ICit^m$			$csaco^m$	388	125	385	26	512	141	319	143

Table S1: (continued.)

	$ICit^m$	→	aKG^m	+	CO_2	*icdh*^{m}	388	125	385	26	512	141	319	143
	aKG^m	→	Scn^m	+	CO_2	*akgdh*^{m}	319	132	313	30	435	148	280	147
	Scn^m	→	Mal^m			$sdh_{f1}{}^m$	319	132	313	30	435	148	280	147
	Scn^m	↔	Mal^m			$sdh_{f2}{}^m$	**270**	**57**	**294**	**30**	**322**	**56**	**94**	**13**
	reversibility %						0.71	0.06	0.98	0.01	0.74	0.19	0.78	0.3
	Mal^m	↔	OAA^m			$mdh_f{}^m$	411	125	408	25	538	142	341	143
	reversibility %						0.49	0.37	0.48	0.36	0.55	0.35	0.51	0.33

Reaction						Reaction	PGM-null 24		PGM-null 38		PGM-KO 24		PGM-KO 38	
						Name	Average		Avearge		Average		Average	
							Net Flux	S.D.	Net Flux	S.D.	Net Flux	S.D.	Net Flux	S.D.
Anaplerotic reactions														
PEP^c	+	CO_2	→	OAA^c			*ppc*^{c}	**293**	**42**	**440**	**12**	**226**	**43**	**212**	**46**
		Mal^m	→	Pyr^m	+	CO_2	*me*^{m}	86	59	109	11	79	40	52	28
		Mal^p	→	Pyr^p	+	CO_2	*me*^{p}	84	38	108	7	98	25	85	15
Substrate entry														
		Suc^{ext}	→	$G6P^c$	+	$F6P^c$	*subs1*	**1013**	**55**	**909**	**23**	**1030**	**28**	**836**	**26**
		Gln^{ext}	↔	Gln^p			$subs2_f$	**465**	**22**	**430**	**6**	**681**	**33**	**509**	**20**
		reversibility %						0.9	0.01	0.76	0.01	0.74	0.19	0.93	0.05
Glutamate assimilation														
		Gln^p	→	Glu^p			as_f	**583**	**93**	**638**	**68**	**418**	**28**	**216**	**39**
		Glu^p	→	aKG^p			$gdh_f{}^p$	**600**	**93**	**653**	**68**	**873**	**46**	**312**	**39**
aKG^p	+	Gln^p	→	Glu^p	+	Glu^p	*gogat*^{p}	296	94	323	61	314	61	173	55
Malate shuttle														
		Mal^m	↔	Mal^c			$malT1_f$	**-107**	**8**	**-205**	**15**	**11**	**7**	**26**	**20**
		reversibility %						0.91	0.09	0.9	0	0.97	0.01	0.95	0.02
		Mal^c	→	Mal^p			*malT2*	221	38	235	14	216	32	211	26

Table S1: (continued.)

Reaction					Reaction Name	PGM-null 24 Average		PGM-null 38 Averge		PGM-KO 24 Average		PGM-KO 38 Average	
						Net Flux	S.D.	Net Flux	S.D.	Net Flux	S.D.	Net Flux	S.D.
Mal^p	→	OAA^p			*mdh^p*	129	8	127	7	142	7	126	10
Biosynthesis of Ser and Gly													
3PG	→	Ser			*bios1*	45	2	46	2	49	2	48	3
Ser	↔	Gly	+	C1	*$bios2_f$*	45	2	46	2	49	2	48	3
reversibility %						0.43	0.04	0.38	0.01	0.39	0.04	0.46	0.03
Thr	→	Gly			*bios3*	35	9	25	2	26	9	19	8
Fluxes towards biosynthesis (other than those related to Ser and Gly) and effluxes into medium													
CO_2	→				*resp*	**3849**	**264**	**3218**	**54**	**4712**	**228**	**3170**	**215**
$G6P^c$	→	biomass			*bios4*	272	15	298	7	281	12	280	12
$G6P^p$	→	biomass			*bios5*	**149**	**6**	**149**	**4**	**4**	**1**	**4**	**1**
$F6P^c$	→	biomass			*bios6*	28	2	29	3	28	3	28	2
$P5P^c$	→	biomass			*bios7*	**208**	**7**	**209**	**3**	**287**	**14**	**276**	**12**
$P5P^p$	→	biomass			*bios8*	9	1	8	0	10	1	8	1
$E4P^p$	→	biomass			*bios9*	21	1	22	2	22	1	21	1
PEP^p	→	biomass			*bios10*	**43**	**3**	**44**	**2**	**37**	**2**	**63**	**2**
ACA^p	→	biomass			*bios11*	**869**	**10**	**788**	**6**	**917**	**10**	**853**	**12**
ACA^p	→	biomass			*bios12*	**21**	**1**	**19**	**2**	**31**	**1**	**27**	**1**
Glycerol	→	biomass			*glyc*	24	13	28	12	17	10	14	10
Pyr	→	biomass			*bios13*	**150**	**7**	**130**	**18**	**165**	**13**	**212**	**23**
OAA	→	biomass			*bios14*	129	8	127	7	124	7	103	10
Glu	→	biomass			*bios15*	**100**	**3**	**90**	**4**	**121**	**9**	**153**	**11**
					Chisquare	100	22	112	33	79	15	76	14

Table S2: Relative multiplet intensities of amino acids from protein and glucosyl units from starch hydrolysates with their standard deviations (SDs) from 2-D [^{13}C, ^{1}H] HSQC spectrum of PGM-null 24 cultures. s indicates singlet, d1 and d2 indicate the first and second doublet and dd indicates the double doublet. Bold-faced carbon atom in Isotopomer indicates labeled carbon atom, normal font indicates unlabeled carbon atom and x indicates unknown labeling state of the carbon atom. Subscripts p, c and m for precursor indicates the compartments plastid, cytosol and mitochondrion in which they are present. RI and SD represent relative intensities and standard deviations respectively.

Hydroly-	Cross peak	PGM-null 24 culture 1		PGM-null 24 culture 2		PGM-null 24 culture 3			
sate	(multiplet)	RI	SD	RI	SD	RI	SD	Precursor	Isotopomer
Protein	Ala α (s)	0.21	0.04	0.20	0.11	0.20	0.08	Pyr	[1**2**3]
Protein	Ala α (d1)	0.09	0.04	0.09	0.11	0.08	0.08	Pyr	[**12**3]
Protein	Ala α (d2)	0.30	0.04	0.11	0.11	0.11	0.08	Pyr	[1**23**]
Protein	Ala α (dd)	0.41	0.04	0.60	0.11	0.62	0.08	Pyr	[**123**]
Protein	Ala β (s)	0.16	0.02	0.31	0.02	0.38	0.05	Pyr	[x2**3**]
Protein	Ala β (d)	0.84	0.02	0.69	0.02	0.62	0.05	Pyr	[x**23**]
Protein	Arg β (s)	0.64	0.07	0.73	0.08	0.69	0.05	Glu_p	[x2**3**4x]
Protein	Arg β (d)	0.27	0.07	0.27	0.08	0.31	0.05	Glu_p	[x**23**4x] + [x2**34**x]
Protein	Arg β (t)	0.18	0.07	0.27	0.08	0.00	0.05	Glu_p	[x**234**x]
Protein	Arg δ (s)	0.54	0.08	0.44	0.10	0.62	0.05	Glu_p	[xxx4**5**]
Protein	Arg δ (d)	0.46	0.08	0.56	0.10	0.38	0.05	Glu_p	[xxx**45**]
Protein	Asp α (s)	0.31	0.02	0.30	0.06	0.34	0.07	OAA	[1**2**3x]
Protein	Asp α (d1)	0.16	0.02	0.15	0.06	0.16	0.07	OAA	[**12**3x]
Protein	Asp α (d2)	0.24	0.02	0.27	0.06	0.25	0.07	OAA	[1**23**x]
Protein	Asp α (dd)	0.29	0.02	0.28	0.06	0.25	0.07	OAA	[**123**x]
Protein	Asp β (s)	0.21	0.03	0.32	0.03	0.31	0.04	OAA	[x2**3**4]
Protein	Asp β (d1)	0.44	0.03	0.35	0.03	0.39	0.04	OAA	[x**23**4]
Protein	Asp β (d2)	0.27	0.03	0.15	0.03	0.13	0.04	OAA	[x2**34**]
Protein	Asp β (dd)	0.09	0.03	0.19	0.03	0.17	0.04	OAA	[x**234**]
Protein	Glu β (s)	0.59	0.06	0.53	0.03	0.50	0.10	Glu	[x2**3**4x]

Table S2: (continued.)

Protein	Glu β (d)	0.47	0.06	0.47	0.03	0.50	0.10	Glu	[x**23**4x] + [x2**34**x]
Protein	Glu β (t)	0.00	0.06	0.00	0.03	0.00	0.10	Glu	[x**234**x]
Protein	Glu γ (s)	0.38	0.05	0.49	0.05	0.49	0.04	Glu	[xx3**45**]
Protein	Glu γ (d1)	0.01	0.05	0.03	0.05	0.01	0.04	Glu	[xx**345**]
Protein	Glu γ (d2)	0.60	0.05	0.46	0.05	0.48	0.04	Glu	[xx3**45**]
Protein	Glu γ (dd)	0.02	0.05	0.02	0.05	0.03	0.04	Glu	[xx**345**]
Protein	Gly α (s)	0.57	0.04	0.60	0.06	0.59	0.59	T3P/GOx	[1**2**x]
Protein	Gly α (d)	0.44	0.04	0.40	0.06	0.41	0.41	T3P/GOx	[**12**x]
Protein	His β (s)	0.79	0.06	0.48	0.07	0.49	0.06	$P5P_p$	[x2**3**4x]
Protein	His β (d1)	0.30	0.06	0.52	0.07	0.51	0.06	$P5P_p$	[x**23**4x]
Protein	His β (d2)	0.69	0.06	0.73	0.07	0.73	0.06	$P5P_p$	[x2**34**x]
Protein	His β (dd)	0.02	0.06	0.09	0.07	0.09	0.06	$P5P_p$	[x**234**x]
Protein	His δ2 (s)	0.19	0.09	0.15	0.09	0.15	0.09	$P5P_p$	[**1**2xxx]
Protein	His δ2 (d)	0.10	0.09	0.03	0.09	0.03	0.09	$P5P_p$	[**12**xxx]
Protein	Ile α (s)	0.47	0.06	0.50	0.09	0.50	0.08	OAA_p/Pyr_p	*[12xx]*·[x2x]
Protein	Ile α (d1)	0.36	0.06	0.32	0.09	0.31	0.08	OAA_p/Pyr_p	*[**12**xx]*·[x2x]
Protein	Ile α (d2)	0.12	0.06	0.12	0.09	0.11	0.08	OAA_p/Pyr_p	*[12xx]*·[x**2**x]
Protein	Ile α (dd)	0.05	0.06	0.06	0.09	0.08	0.08	OAA_p/Pyr_p	*[**12**xx]*·[x**2**x]
Protein	Ile γ1(s)	0.65	0.06	0.62	0.09	0.53	0.10	Pyr_p/OAA_p	[x2x]·*[xx**3**4]*
Protein	Ile γ1(d)	0.35	0.06	0.38	0.09	0.47	0.10	Pyr_p/OAA_p	[x**2**x]·*[xx**3**4]* + [x2x]·*[xx**34**]*
Protein	Ile γ1(t)	0.00	0.06	0.00	0.09	0.00	0.10	Pyr_p/OAA_p	[x**2**x]·*[xx**34**]*

Hydroly-	**Cross peak**	**PGM-null 24 culture 1**		**PGM-null 24 culture 2**		**PGM-null 24 culture 3**			
sate	**(multiplet)**	**RI**	**SD**	**RI**	**SD**	**RI**	**SD**	**Precursor**	**Isotopomer**

Table S2: (continued.)

Protein	Ile γ2 (s)	0.74	0.03	0.64	0.03	0.79	0.05	Pyr_p	[x2**3**]
Protein	Ile γ2 (d)	0.26	0.03	0.36	0.03	0.21	0.05	Pyr_p	[x**23**]
Protein	Ile δ (s)	0.63	0.06	0.59	0.08	0.66	0.11	OAA_p	[xx3**4**]
Protein	Ile δ (d)	0.37	0.06	0.41	0.08	0.34	0.11	OAA_p	[xx**34**]
Protein	Leu α (s)	0.45	0.05	0.26	0.08	0.26	0.10	$ACoA_p$/Pyr_p	*[**12**]*·[x2x]
Protein	Leu α (d1)	0.47	0.05	0.66	0.08	0.66	0.10	$ACoA_p$/Pyr_p	*[**12**]*·[x2x]
Protein	Leu α (d2)	0.03	0.05	0.03	0.08	0.03	0.10	$ACoA_p$/Pyr_p	*[12]*·[x**2**x]
Protein	Leu α (dd)	0.06	0.05	0.05	0.08	0.55	0.10	$ACoA_p$/Pyr_p	*[**12**]*·[x**2**x]
Protein	Leu b (s)	0.86	0.05	0.88	0.05	0.79	0.05	$ACoA_p$/Pyr_p	*[x2]*·[x**2**x].[x2x]
Protein	Leu b (d)	0.14	0.05	0.13	0.05	0.21	0.05	$ACoA_p$/Pyr_p	*[x**2**]*·[x**2**x].[x2x]+*[x2]·[x2x].[x2x]*
Protein	Leu b (t)	0.00	0.05	0.00	0.05	0.00	0.05	$ACoA_p$/Pyr_p	*[x**2**]*·[x**2**x].[x**2**x]
Protein	Leu δ1 (s)	0.39	0.03	0.39	0.04	0.42	0.04	Pyr_p	[x2**3**]
Protein	Leu δ1 (d)	0.61	0.03	0.61	0.04	0.58	0.04	Pyr_p	[x**23**]
Protein	Leu δ2 (s)	0.90	0.02	0.89	0.02	0.81	0.05	Pyr_p	[x2x]·[xx**3**]
Protein	Leu δ2 (d)	0.10	0.02	0.11	0.02	0.19	0.05	Pyr_p	[x**2**x]·[xx**3**]
Protein	Lys β (s)	0.18	0.06	0.19	0.11	0.19	0.06	OAA_p/Pyr_p	½{*[x2**3**4]* + [x2**3**]·*[xxx4]*}
Protein	Lys β (d)	0.82	0.06	0.81	0.11	0.81	0.06	OAA_p/Pyr_p	½{*[x**23**4]* + *[x2**34**]* +
								OAA_p/Pyr_p	[x**23**]·*[xxx4]* + [x2**3**]·*[xxx**4**]*}
Protein	Lys β (t)	0.00	0.06	0.00	0.11	0.00	0.06	OAA_p/Pyr_p	½{*[x**234**]* + [x**23**]·*[xxx**4**]*}
Protein	Lys γ (s)	0.63	0.05	0.52	0.10	0.58	0.08	OAA_p/Pyr_p	*[xx3**4**]*·[xx3]
Protein	Lys γ (d)	0.37	0.05	0.44	0.10	0.42	0.08	OAA_p/Pyr_p	*[xx**34**]*·[xx3] + *[xx3**4**]·[xx**3**]*
Protein	Lys γ (t)	0.00	0.05	0.00	0.10	0.00	0.08	OAA_p/Pyr_p	*[xx**34**]*·[xx**3**]
Protein	Lys δ (s)	0.42	0.06	0.35	0.07	0.40	0.08	OAA_p/Pyr_p	½{*[x2**3**4]* + [x2**3**]·*[xxx4]*}
Protein	Lys δ (d)	0.58	0.06	0.65	0.07	0.60	0.08	OAA_p/Pyr_p	½{*[x**23**4]* + *[x2**34**]* +

Table S2: (continued.)

									[x**23**]·*[xxx4]* + [x2**3**]·*[xxx**4**]*}
Protein	Lys δ (t)	0.00	0.06	0.00	0.07	0.00	0.08	OAA_p/Pyr_p	½{*[x**234**]* + [x**23**]·*[xxx**4**]*}
Protein	Lys ε (s)	0.41	0.06	0.33	0.06	0.36	0.07	OAA_p/Pyr_p	½{[x**2**3] + *[x**2**3x]*}
Protein	Lys ε (d)	0.59	0.06	0.67	0.06	0.64	0.07	OAA_p/Pyr_p	½{[x**23**] + *[x**23**x]*}
Protein	Phe α (s)	0.22	0.10	0.34	0.10	0.22	0.08	PEP_p	[1**2**3]
Protein	Phe α (d1)	0.05	0.10	0.04	0.10	0.05	0.08	PEP_p	[**12**3]
Protein	Phe α (d2)	0.11	0.10	0.49	0.10	0.60	0.08	PEP_p	[1**23**]
Protein	Phe α (dd)	0.12	0.10	0.12	0.10	0.14	0.08	PEP_p	[**123**]
Protein	Phe β (s)	0.54	0.06	0.34	0.08	0.32	0.05	PEP_p	[x2**3**]·[x2x]
Protein	Phe β (d1)	0.35	0.06	0.55	0.08	0.57	0.05	PEP_p	[x**23**]·[x2x]
Protein	Phe β (d2)	0.07	0.06	0.04	0.08	0.04	0.05	PEP_p	[x2**3**]·[x**2**x]
Protein	Phe β (dd)	0.04	0.06	0.07	0.08	0.07	0.05	PEP_p	[x**23**]·[x**2**x]
Protein	Pro γ (s)	0.48	0.07	0.45	0.08	0.44	0.08	Glu_p	[xx3**4**5]
Protein	Pro γ (d)	0.52	0.07	0.55	0.08	0.56	0.08	Glu_p	[xx**34**5] + [xx3**45**]
Protein	Pro γ (t)	0.00	0.07	0.00	0.08	0.00	0.08	Glu_p	[xx**345**]
Protein	Pro δ (s)	0.44	0.04	0.36	0.11	0.42	0.05	Glu_p	[xxx4**5**]
Protein	Pro δ (d)	0.56	0.04	0.64	0.11	0.58	0.05	Glu_p	[xxx**45**]

Hydroly-	**Cross peak**	**PGM-null 24 culture 1**		**PGM-null 24 culture 2**		**PGM-null 24 culture 3**			
sate	**(multiplet)**	**RI**	**SD**	**RI**	**SD**	**RI**	**SD**	**Precursor**	**Isotopomer**
Protein	Ser α (s)	0.24	0.09	0.24	0.09	0.25	0.07	Ser	[1**2**3]
Protein	Ser α (d1)	0.20	0.09	0.16	0.09	0.21	0.07	Ser	[**12**3]
Protein	Ser α (d2)	0.09	0.09	0.16	0.09	0.12	0.07	Ser	[1**23**]
Protein	Ser α (dd)	0.47	0.09	0.44	0.09	0.42	0.07	Ser	[**123**]

Table S2: (continued.)

Protein	Ser β (s)	0.46	0.07	0.56	0.06	0.56	0.06	Ser	[x2**3**]
Protein	Ser β (d)	0.54	0.07	0.45	0.06	0.44	0.06	Ser	[x**23**]
Protein	Thr α (s)	0.35	0.09	0.33	0.09	0.34	0.07	OAA_p	[1**2**3x]
Protein	Thr α (d1)	0.14	0.09	0.15	0.09	0.14	0.07	OAA_p	[**12**3x]
Protein	Thr α (d2)	0.32	0.09	0.28	0.09	0.29	0.07	OAA_p	[1**23**x]
Protein	Thr α (dd)	0.19	0.09	0.24	0.09	0.23	0.07	OAA_p	[**123**x]
Protein	Thr γ2 (s)	0.57	0.03	0.64	0.02	0.73	0.03	OAA_p	[xx3**4**]
Protein	Thr γ2 (d)	0.43	0.03	0.36	0.02	0.27	0.03	OAA_p	[xx**34**]
Protein	Tyr α (s)	0.54	0.10	0.20	0.08	0.19	0.10	PEP_p	[1**2**3]
Protein	Tyr α (d1)	0.08	0.10	0.09	0.08	0.06	0.10	PEP_p	[**12**3]
Protein	Tyr α (d2)	0.09	0.10	0.49	0.08	0.58	0.10	PEP_p	[1**23**]
Protein	Tyr α (dd)	0.28	0.10	0.12	0.08	0.17	0.10	PEP_p	[**123**]
Protein	Tyr β (s)	0.54	0.05	0.12	0.10	0.31	0.05	PEP_p	[x2**3**]·[x2x]
Protein	Tyr β (d)	0.41	0.05	0.78	0.10	0.64	0.05	PEP_p	[x**23**]·[x2x] + [x23]·[x2x]
Protein	Tyr β (t)	0.04	0.05	0.20	0.10	0.10	0.05	PEP_p	[x**23**]·[x**2**x]
Protein	Tyr δ1 (s)	0.53	0.04	0.27	0.07	0.49	0.08	PEP_p/$E4P_p$	*[x23]*·[1xxx]
Protein	Tyr δ1 (d)	0.42	0.04	0.07	0.07	0.51	0.08	PEP_p/$E4P_p$	*[x**23**]*·[1xxx] + *[x23]*·[**1**xxx]
Protein	Tyr δ1 (t)	0.04	0.04	0.00	0.07	0.00	0.08	PEP_p/$E4P_p$	*[x**23**]*·[**1**xxx]
Protein	Tyr ε1 (s)	0.61	0.07	0.68	0.04	0.55	0.06	PEP_p/$E4P_p$	*[xx3]*·[**1**2xx]
Protein	Tyr ε1 (d)	0.17	0.07	0.37	0.04	0.37	0.06	PEP_p/$E4P_p$	*[xx**3**]*·[**1**2xx] + *[xx3]*·[**12**xx]
Protein	Tyr ε1 (t)	0.44	0.07	0.07	0.04	0.07	0.06	PEP_p/$E4P_p$	*[xx**3**]*·[**12**xx]
Protein	Val α (s)	0.36	0.07	0.73	0.10	0.69	0.06	Pyr_p	[1**2**x]·[x2x]
Protein	Val α (d1)	0.53	0.07	0.41	0.10	0.21	0.06	Pyr_p	[**12**x]·[x2x]
Protein	Val α (d2)	0.03	0.07	0.02	0.10	0.03	0.06	Pyr_p	[1**2**x]·[x**2**x]

Table S2: (continued.)

Protein	Val α (dd)	0.08	0.07	0.26	0.10	0.11	0.06	Pyr_p	[1**2**x]·[x**2**x]
Protein	Val γ1 (s)	0.49	0.02	0.39	0.03	0.30	0.03	Pyr_p	[x2**3**]
Protein	Val γ1 (d)	0.51	0.02	0.61	0.03	0.71	0.03	Pyr_p	[x**23**]
Protein	Val γ2 (s)	0.95	0.02	0.93	0.02	0.90	0.07	Pyr_p	[x2x]·[xx**3**]
Protein	Val γ2 (d)	0.05	0.02	0.07	0.02	0.10	0.07	Pyr_p	[x**2**x]·[xx**3**]

Hydroly-	**Cross peak**	**PGM-null 24 culture 1**		**PGM-null 24 culture 2**		**PGM-null 24 culture 3**			
sate	**(multiplet)**	**RI**	**SD**	**RI**	**SD**	**RI**	**SD**	**Precursor**	**Isotopomer**
Protein	HMF 1 (s)	0.44	0.09	0.45	0.09	0.15	0.05	$G6P_c/F6P_c$	[**1**2xxxx]
Protein	HMF 1 (d)	0.56	0.09	0.55	0.09	0.85	0.05	$G6P_c/F6P_c$	[**12**xxxx]
Protein	LVA 3 (s)	0.66	0.04	0.65	0.03	0.69	0.04	$G6P_c/F6P_c$	[x2**3**4xx]
Protein	LVA 3 (d1)	0.13	0.04	0.11	0.03	0.16	0.04	$G6P_c/F6P_c$	[x**23**4xx]
Protein	LVA 3 (d2)	0.12	0.04	0.11	0.03	0.12	0.04	$G6P_c/F6P_c$	[x2**34**xx]
Protein	LVA 3 (dd)	0.09	0.04	0.13	0.03	0.03	0.04	$G6P_c/F6P_c$	[x**234**xx]
Protein	LVA 6 (s)	0.18	0.02	0.21	0.03	0.18	0.05	$G6P_c/F6P_c$	[xxx45**6**]
Protein	LVA 6 (d1)	0.00	0.02	0.02	0.03	0.01	0.05	$G6P_c/F6P_c$	[xxx4**56**]
Protein	LVA 6 (d2)	0.31	0.02	0.25	0.03	0.30	0.05	$G6P_c/F6P_c$	[xxx**45**6]
Protein	LVA 6 (dd)	0.50	0.02	0.52	0.03	0.51	0.05	$G6P_c/F6P_c$	[xxx**456**]
Starch	HMF 1 (s)	0.65	0.04	0.42	0.07	0.27	0.05	$G6P_p$	[**1**2xxxx]
Starch	HMF 1 (d)	0.35	0.04	0.58	0.07	0.73	0.05	$G6P_p$	[**12**xxxx]
Starch	LVA 3 (s)	0.72	0.01	0.65	0.02	0.62	0.02	$G6P_p$	[x2**3**4xx]
Starch	LVA 3 (d1)	0.15	0.01	0.10	0.02	0.10	0.02	$G6P_p$	[x**23**4xx]
Starch	LVA 3 (d2)	0.09	0.01	0.08	0.02	0.08	0.02	$G6P_p$	[x2**34**xx]

Table S2: (contined.)

Starch	LVA 3 (dd)	0.04	0.01	0.16	0.02	0.21	0.02	$G6P_p$	[x**234**xx]
Starch	LVA 6 (s)	0.53	0.01	0.29	0.01	0.27	0.02	$G6P_p$	[xxx45**6**]
Starch	LVA 6 (d1)	0.06	0.01	0.03	0.01	0.03	0.02	$G6P_p$	[xxx4**56**]
Starch	LVA 6 (d2)	0.26	0.01	0.41	0.01	0.40	0.02	$G6P_p$	[xxx**4**5**6**]
Starch	LVA 6 (dd)	0.15	0.01	0.26	0.01	0.30	0.02	$G6P_p$	[xxx**456**]

Table S3: Relative multiplet intensities of amino acids from protein and glucosyl units from starch hydrolysates with their standard deviations (SDs) from 2-D [^{13}C, ^{1}H] HSQC spectrum of PGM-null 38 cultures. s indicates singlet, d1 and d2 indicate the first and second doublet and dd indicates the double doublet. Bold-faced carbon atom in Isotopomer indicates labeled carbon atom, normal font indicates unlabeled carbon atom and x indicates unknown labeling state of the carbon atom. Subscripts p, c and m for precursor indicates the compartments plastid, cytosol and mitochondrion in which they are present. RI and SD represent relative intensities and standard deviations respectively.

Hydroly-sate	Cross peak (multiplet)	PGM-null 38 culture 1		PGM-null 38 culture 2		PGM-null 38 culture 3			
		RI	SD	RI	SD	RI	SD	Precursor	Isotopomer
Protein	Ala α (s)	0.20	0.06	0.16	0.06	0.22	0.07	Pyr	[1**2**3]
Protein	Ala α (d1)	0.10	0.06	0.03	0.06	0.08	0.07	Pyr	[**12**3]
Protein	Ala α (d2)	0.34	0.06	0.49	0.06	0.12	0.07	Pyr	[1**23**]
Protein	Ala α (dd)	0.37	0.06	0.32	0.06	0.58	0.07	Pyr	[**123**]
Protein	Ala β (s)	0.18	0.02	0.22	0.01	0.17	0.02	Pyr	[x2**3**]
Protein	Ala β (d)	0.82	0.02	0.78	0.01	0.83	0.02	Pyr	[x**23**]
Protein	Arg β (s)	0.73	0.06	0.73	0.08	0.70	0.07	Glu_p	[x2**3**4x]
Protein	Arg β (d)	0.27	0.06	0.27	0.08	0.30	0.07	Glu_p	[x**23**4x] + [x2**34**x]
Protein	Arg β (t)	0.00	0.06	0.00	0.08	0.00	0.07	Glu_p	[x**234**x]
Protein	Arg δ (s)	0.63	0.08	0.40	0.09	0.42	0.09	Glu_p	[xxx4**5**]
Protein	Arg δ (d)	0.37	0.08	0.60	0.09	0.58	0.09	Glu_p	[xxx**45**]
Protein	Asp α (s)	0.29	0.08	0.30	0.08	0.31	0.04	OAA	[1**2**3x]
Protein	Asp α (d1)	0.15	0.08	0.13	0.08	0.14	0.04	OAA	[**12**3x]
Protein	Asp α (d2)	0.27	0.08	0.31	0.08	0.27	0.04	OAA	[1**23**x]
Protein	Asp α (dd)	0.29	0.08	0.25	0.08	0.28	0.04	OAA	[**123**x]
Protein	Asp β (s)	0.21	0.03	0.24	0.02	0.33	0.04	OAA	[x2**3**4]
Protein	Asp β (d1)	0.35	0.03	0.30	0.02	0.38	0.04	OAA	[x**23**4]
Protein	Asp β (d2)	0.37	0.03	0.41	0.02	0.12	0.04	OAA	[x2**34**]
Protein	Asp β (dd)	0.07	0.03	0.05	0.02	0.17	0.04	OAA	[x**234**]

Table S3: (continued.)

Protein	Glu β (s)	0.64	0.04	0.66	0.06	0.51	0.08	Glu	[x23**4**x]
Protein	Glu β (d)	0.33	0.04	0.34	0.06	0.49	0.08	Glu	[x**23**4x] + [x2**34**x]
Protein	Glu β (t)	0.03	0.04	0.00	0.06	0.00	0.08	Glu	[x**234**x]
Protein	Glu γ (s)	0.44	0.03	0.46	0.02	0.50	0.03	Glu	[xx3**4**5]
Protein	Glu γ (d1)	0.01	0.03	0.02	0.02	0.03	0.03	Glu	[xx**34**5]
Protein	Glu γ (d2)	0.51	0.03	0.50	0.02	0.44	0.03	Glu	[xx3**45**]
Protein	Glu γ (dd)	0.03	0.03	0.02	0.02	0.03	0.03	Glu	[xx**345**]
Protein	Gly α (s)	0.53	0.05	0.61	0.03	0.59	0.05	T3P/GOx	[1**2**x]
Protein	Gly α (d)	0.47	0.05	0.39	0.03	0.41	0.05	T3P/GOx	[**12**x]
Protein	His β (s)	0.61	0.08	0.51	0.05	0.42	0.09	$P5P_p$	[x23**4**x]
Protein	His β (d1)	0.39	0.08	0.49	0.05	0.58	0.09	$P5P_p$	[x**23**4x]
Protein	His β (d2)	0.30	0.08	0.66	0.05	0.58	0.09	$P5P_p$	[x2**34**x]
Protein	His β (dd)	0.00	0.08	0.08	0.05	0.07	0.09	$P5P_p$	[x**234**x]
Protein	His δ2 (s)	0.59	0.07	0.21	0.08	0.30	0.09	$P5P_p$	[**1**2xxx]
Protein	His δ2 (d)	0.10	0.07	0.05	0.08	0.05	0.09	$P5P_p$	[**12**xxx]
Protein	Ile α (s)	0.54	0.06	0.54	0.06	0.49	0.08	OAA_p/Pyr_p	*[**12**xx]*·[x2x]
Protein	Ile α (d1)	0.33	0.06	0.33	0.06	0.33	0.08	OAA_p/Pyr_p	*[**12**xx]*·[x2x]
Protein	Ile α (d2)	0.12	0.06	0.04	0.06	0.04	0.08	OAA_p/Pyr_p	*[**12**xx]*·[x**2**x]
Protein	Ile α (dd)	0.02	0.06	0.09	0.06	0.14	0.08	OAA_p/Pyr_p	*[**12**xx]*·[x**2**x]

Hydroly-	**Cross peak**	**PGM-null 38 culture 1**		**PGM-null 38 culture 2**		**PGM-null 38 culture 3**			
sate	**(multiplet)**	**RI**	**SD**	**RI**	**SD**	**RI**	**SD**	**Precursor**	**Isotopomer**
Protein	Ile γ1(s)	0.68	0.08	0.67	0.08	0.65	0.09	Pyr_p/OAA_p	[x2x]·*[xx**34**]*
Protein	Ile γ1(d)	0.32	0.08	0.33	0.08	0.35	0.09	Pyr_p/OAA_p	[x**2**x]·*[xx**34**]* + [x2x]·*[xx**34**]*
Protein	Ile γ1(t)	0.00	0.08	0.00	0.08	0.00	0.09	Pyr_p/OAA_p	[x**2**x]·*[xx**34**]*

Table S3: (continued.)

Protein	Ile γ2 (s)	0.53	0.03	0.48	0.02	0.65	0.04	Pyr_p	[x2**3**]
Protein	Ile γ2 (d)	0.47	0.03	0.52	0.02	0.35	0.04	Pyr_p	[x**23**]
Protein	Ile δ (s)	0.46	0.05	0.34	0.04	0.36	0.06	OAA_p	[xx3**4**]
Protein	Ile δ (d)	0.54	0.05	0.66	0.04	0.64	0.06	OAA_p	[xx**34**]
Protein	Leu α (s)	0.25	0.07	0.26	0.08	0.26	0.05	$ACoA_p/Pyr_p$	*[**12**]*·[x2x]
Protein	Leu α (d1)	0.63	0.07	0.65	0.08	0.65	0.05	$ACoA_p/Pyr_p$	*[**12**]*·[x2x]
Protein	Leu α (d2)	0.03	0.07	0.03	0.08	0.04	0.05	$ACoA_p/Pyr_p$	*[**12**]*·[x**2**x]
Protein	Leu α (dd)	0.10	0.07	0.06	0.08	0.05	0.05	$ACoA_p/Pyr_p$	*[**12**]*·[x**2**x]
Protein	Leu b (s)	0.73	0.07	0.90	0.04	0.93	0.04	$ACoA_p/Pyr_p$	*[x2]*·[x**2**x].[x2x]
Protein	Leu b (d)	0.27	0.07	0.10	0.04	0.07	0.04	$ACoA_p/Pyr_p$	*[x**2**]*·[x**2**x].[x2x]+*[x2]·[x2x].[x2x]*
Protein	Leu b (t)	0.00	0.07	0.00	0.04	0.00	0.04	$ACoA_p/Pyr_p$	*[x**2**]*·[x**2**x].[x**2**x]
Protein	Leu δ1 (s)	0.34	0.06	0.39	0.02	0.28	0.03	Pyr_p	[x2**3**]
Protein	Leu δ1 (d)	0.66	0.06	0.61	0.02	0.72	0.03	Pyr_p	[x**23**]
Protein	Leu δ2 (s)	0.91	0.04	0.91	0.02	0.88	0.04	Pyr_p	[x2x]·[xx**3**]
Protein	Leu δ2 (d)	0.09	0.04	0.09	0.02	0.12	0.04	Pyr_p	[x**2**x]·[xx**3**]
Protein	Lys β (s)	0.23	0.07	0.35	0.06	0.17	0.07	OAA_p/Pyr_p	½{*[x2**34**]* + [x2**3**]·*[xxx4]*}
Protein	Lys β (d)	0.68	0.07	0.59	0.06	0.73	0.07	OAA_p/Pyr_p	½{*[x**23**4]* + *[x2**34**]* +
								OAA_p/Pyr_p	[x**23**]·*[xxx4]* + [x2**3**]·*[xxx**4**]*}
Protein	Lys β (t)	0.17	0.07	0.12	0.06	0.19	0.07	OAA_p/Pyr_p	½{*[x**234**]* + [x**23**]·*[xxx**4**]*}
Protein	Lys γ (s)	0.37	0.09	0.43	0.09	0.62	0.07	OAA_p/Pyr_p	*[xx3**4**]*·[xx3]
Protein	Lys γ (d)	0.63	0.09	0.52	0.09	0.38	0.07	OAA_p/Pyr_p	*[xx**34**]*·[xx3] + *[xx3**4**]·[xx**3**]*
Protein	Lys γ (t)	0.00	0.09	0.06	0.09	0.00	0.07	OAA_p/Pyr_p	*[xx**34**]*·[xx**3**]
Protein	Lys δ (s)	0.30	0.03	0.35	0.06	0.21	0.09	OAA_p/Pyr_p	½{*[x2**34**]* + [x2**3**]·*[xxx4]*}
Protein	Lys δ (d)	0.62	0.03	0.58	0.06	0.67	0.09	OAA_p/Pyr_p	½{*[x**23**4]* + *[x2**34**]* +

Table S3: (continued.)

									[x**23**]·*[xxx4]* + [x2**3**]·*[xxx**4**]*}
Protein	Lys δ (t)	0.15	0.03	0.15	0.06	0.23	0.09	OAA_p/Pyr_p	½{*[x**234**]* + [x**23**]·*[xxx**4**]*}
Protein	Lys ε (s)	0.41	0.03	0.38	0.03	0.41	0.05	OAA_p/Pyr_p	½{[x**2**3] + *[x**2**3x]*}
Protein	Lys ε (d)	0.59	0.03	0.62	0.03	0.59	0.05	OAA_p/Pyr_p	½{[x**23**] + *[x**23**x]*}
Protein	Phe α (s)	0.20	0.12	0.34	0.08	0.21	0.08	PEP_p	[1**2**3]
Protein	Phe α (d1)	0.01	0.12	0.04	0.08	0.04	0.08	PEP_p	[**12**3]
Protein	Phe α (d2)	0.13	0.12	0.45	0.08	0.11	0.08	PEP_p	[1**23**]
Protein	Phe α (dd)	0.65	0.12	0.16	0.08	0.64	0.08	PEP_p	[**123**]
Protein	Phe β (s)	0.30	0.06	0.34	0.08	0.21	0.09	PEP_p	[x2**3**]·[x2x]
Protein	Phe β (d1)	0.59	0.06	0.55	0.08	0.06	0.09	PEP_p	[x**23**]·[x2x]
Protein	Phe β (d2)	0.04	0.06	0.04	0.08	0.70	0.09	PEP_p	[x2**3**]·[x**2**x]
Protein	Phe β (dd)	0.07	0.06	0.07	0.08	0.03	0.09	PEP_p	[x**23**]·[x**2**x]
Protein	Pro γ (s)	0.48	0.08	0.44	0.09	0.33	0.08	Glu_p	[xx3**4**5]
Protein	Pro γ (d)	0.52	0.08	0.56	0.09	0.67	0.08	Glu_p	[xx**34**5] + [xx3**45**]
Protein	Pro γ (t)	0.00	0.08	0.00	0.09	0.00	0.08	Glu_p	[xx**345**]
Protein	Pro δ (s)	0.28	0.07	0.61	0.07	0.44	0.08	Glu_p	[xxx4**5**]
Protein	Pro δ (d)	0.72	0.07	0.40	0.07	0.56	0.08	Glu_p	[xxx**45**]

Hydroly-	**Cross peak**	**PGM-null 38 culture 1**		**PGM-null 38 culture 2**		**PGM-null 38 culture 3**			
sate	**(multiplet)**	**RI**	**SD**	**RI**	**SD**	**RI**	**SD**	**Precursor**	**Isotopomer**
Protein	Ser α (s)	0.23	0.09	0.18	0.07	0.24	0.07	Ser	[1**2**3]
Protein	Ser α (d1)	0.18	0.09	0.05	0.07	0.17	0.07	Ser	[**12**3]
Protein	Ser α (d2)	0.18	0.09	0.47	0.07	0.14	0.07	Ser	[1**23**]
Protein	Ser α (dd)	0.42	0.09	0.31	0.07	0.45	0.07	Ser	[**123**]

Table S3: (continued.)

Protein	Ser β (s)	0.50	0.06	0.07	0.09	0.15	0.07	Ser	[x2**3**]
Protein	Ser β (d)	0.50	0.06	0.93	0.09	0.85	0.07	Ser	[x**23**]
Protein	Thr α (s)	0.34	0.07	0.33	0.08	0.30	0.08	OAA_p	[1**2**3x]
Protein	Thr α (d1)	0.13	0.07	0.14	0.08	0.14	0.08	OAA_p	[**12**3x]
Protein	Thr α (d2)	0.29	0.07	0.26	0.08	0.29	0.08	OAA_p	[1**23**x]
Protein	Thr α (dd)	0.23	0.07	0.26	0.08	0.26	0.08	OAA_p	[**123**x]
Protein	Thr γ2 (s)	0.53	0.03	0.48	0.05	0.72	0.02	OAA_p	[xx3**4**]
Protein	Thr γ2 (d)	0.47	0.03	0.52	0.05	0.28	0.02	OAA_p	[xx**34**]
Protein	Tyr α (s)	0.30	0.06	0.34	0.08	0.19	0.10	PEP_p	[1**2**3]
Protein	Tyr α (d1)	0.04	0.06	0.04	0.08	0.05	0.10	PEP_p	[**12**3]
Protein	Tyr α (d2)	0.41	0.06	0.45	0.08	0.17	0.10	PEP_p	[1**23**]
Protein	Tyr α (dd)	0.25	0.06	0.16	0.08	0.59	0.10	PEP_p	[**123**]
Protein	Tyr β (s)	0.16	0.07	0.11	0.10	0.21	0.07	PEP_p	[x2**3**]·[x2x]
Protein	Tyr β (d)	0.75	0.07	0.89	0.10	0.70	0.07	PEP_p	[x**23**]·[x2x] + [x23]·[x2x]
Protein	Tyr β (t)	0.17	0.07	0.00	0.10	0.18	0.07	PEP_p	[x**23**]·[x**2**x]
Protein	Tyr δ1 (s)	0.29	0.07	0.32	0.06	0.30	0.04	PEP_p/$E4P_p$	*[x23]*·[1xxx]
Protein	Tyr δ1 (d)	0.71	0.07	0.68	0.06	0.70	0.04	PEP_p/$E4P_p$	*[x**23**]*·[1xxx] + *[x2**3**]*·[**1**xxx]
Protein	Tyr δ1 (t)	0.00	0.07	0.00	0.06	0.00	0.04	PEP_p/$E4P_p$	*[x**23**]*·[**1**xxx]
Protein	Tyr ε1 (s)	0.60	0.07	0.54	0.05	0.70	0.04	PEP_p/$E4P_p$	*[xx3]*·[**1**2xx]
Protein	Tyr ε1 (d)	0.20	0.07	0.35	0.05	0.15	0.04	PEP_p/$E4P_p$	*[xx**3**]*·[**1**2xx] + *[xx3]*·[**12**xx]
Protein	Tyr ε1 (t)	0.40	0.07	0.11	0.05	0.31	0.04	PEP_p/$E4P_p$	*[xx**3**]*·[**12**xx]
Protein	Val α (s)	0.63	0.07	0.70	0.04	0.37	0.08	Pyr_p	[1**2**x]·[x2x]
Protein	Val α (d1)	0.27	0.07	0.19	0.04	0.51	0.08	Pyr_p	[**12**x]·[x2x]
Protein	Val α (d2)	0.08	0.07	0.09	0.04	0.04	0.08	Pyr_p	[1**2**x]·[x**2**x]

Table S3: (continued.)

Protein	Val α (dd)	0.03	0.07	0.02	0.04	0.09	0.08	Pyr_p	[1**2**x]·[x**2**x]
Protein	Val γ1 (s)	0.35	0.01	0.39	0.01	0.28	0.02	Pyr_p	[x23**]**
Protein	Val γ1 (d)	0.65	0.01	0.61	0.01	0.72	0.02	Pyr_p	[x**23**]
Protein	Val γ2 (s)	0.91	0.02	0.87	0.03	0.94	0.01	Pyr_p	[x2x]·[xx**3**]
Protein	Val γ2 (d)	0.09	0.02	0.13	0.03	0.06	0.01	Pyr_p	[x**2**x]·[xx**3**]

Hydroly-	**Cross peak**	**PGM-null 38 culture 1**		**PGM-null 38 culture 2**		**PGM-null 38 culture 3**			
sate	**(multiplet)**	**RI**	**SD**	**RI**	**SD**	**RI**	**SD**	**Precursor**	**Isotopomer**
Protein	HMF 1 (s)	0.41	0.10	0.14	0.04	0.35	0.10	$G6P_c/F6P_c$	[**1**2xxxx]
Protein	HMF 1 (d)	0.59	0.10	0.86	0.04	0.65	0.10	$G6P_c/F6P_c$	[**12**xxxx]
Protein	LVA 3 (s)	0.54	0.03	0.62	0.05	0.73	0.04	$G6P_c/F6P_c$	[x2**3**4xx]
Protein	LVA 3 (d1)	0.34	0.03	0.24	0.05	0.17	0.04	$G6P_c/F6P_c$	[x**23**4xx]
Protein	LVA 3 (d2)	0.07	0.03	0.09	0.05	0.10	0.04	$G6P_c/F6P_c$	[x2**34**xx]
Protein	LVA 3 (dd)	0.05	0.03	0.05	0.05	0.00	0.04	$G6P_c/F6P_c$	[x**234**xx]
Protein	LVA 6 (s)	0.17	0.03	0.13	0.03	0.14	0.05	$G6P_c/F6P_c$	[xxx45**6**]
Protein	LVA 6 (d1)	0.02	0.03	0.01	0.03	0.01	0.05	$G6P_c/F6P_c$	[xxx4**56**]
Protein	LVA 6 (d2)	0.41	0.03	0.50	0.03	0.35	0.05	$G6P_c/F6P_c$	[xxx**4**5**6**]
Protein	LVA 6 (dd)	0.41	0.03	0.36	0.03	0.50	0.05	$G6P_c/F6P_c$	[xxx**456**]
Starch	HMF 1 (s)	0.32	0.10	0.37	0.13	0.34	0.13	$G6P_p$	[**1**2xxxx]
Starch	HMF 1 (d)	0.68	0.01	0.63	0.13	0.66	0.13	$G6P_p$	[**12**xxxx]
Starch	LVA 3 (s)	0.64	0.01	0.72	0.01	0.52	0.01	$G6P_p$	[x2**3**4xx]
Starch	LVA 3 (d1)	0.25	0.01	0.17	0.01	0.44	0.01	$G6P_p$	[x**23**4xx]
Starch	LVA 3 (d2)	0.08	0.01	0.09	0.01	0.00	0.01	$G6P_p$	[x2**34**xx]
Starch	LVA 3 (dd)	0.03	0.01	0.02	0.01	0.04	0.01	$G6P_p$	[x**234**xx]

Table S3: (continued.)

Starch	LVA 6 (s)	0.30	0.01	0.34	0.01	0.45	0.01	$G6P_p$	[xxx45**6**]
Starch	LVA 6 (d1)	0.03	0.01	0.04	0.01	0.05	0.01	$G6P_p$	[xxx4**56**]
Starch	LVA 6 (d2)	0.41	0.01	0.45	0.01	0.19	0.01	$G6P_p$	[xxx**4**5**6**]
Starch	LVA 6 (dd)	0.25	0.01	0.17	0.01	0.30	0.01	$G6P_p$	[xxx**456**]

Table S4: Relative multiplet intensities of amino acids from protein and glucosyl units from starch hydrolysates with their standard deviations (SDs) from 2-D [^{13}C, ^{1}H] HSQC spectrum of PGM-KO 24 cultures. s indicates singlet, d1 and d2 indicate the first and second doublet and dd indicates the double doublet. Bold-faced carbon atom in Isotopomer indicates labeled carbon atom, normal font indicates unlabeled carbon atom and x indicates unknown labeling state of the carbon atom. Subscripts p, c and m for precursor indicates the compartments plastid, cytosol and mitochondrion in which they are present. RI and SD represent relative intensities and standard deviations respectively.

Hydroly-	Cross peak	PGM-KO 24 culture 1		PGM-KO 24 culture 2		PGM-KO 24 culture 3			
sate	(multiplet)	RI	SD	RI	SD	RI	SD	Precursor	Isotopomer
Protein	Ala α (s)	0.22	0.01	0.23	0.08	0.20	0.07	Pyr	[1**2**3]
Protein	Ala α (d1)	0.10	0.01	0.07	0.08	0.05	0.07	Pyr	[**12**3]
Protein	Ala α (d2)	0.12	0.01	0.20	0.08	0.22	0.07	Pyr	[1**23**]
Protein	Ala α (dd)	0.55	0.01	0.49	0.08	0.53	0.07	Pyr	[**123**]
Protein	Ala β (s)	0.29	0.01	0.19	0.02	0.27	0.03	Pyr	[x2**3**]
Protein	Ala β (d)	0.71	0.01	0.81	0.02	0.73	0.03	Pyr	[x**23**]
Protein	Arg β (s)	0.72	0.02	0.58	0.08	0.72	0.08	Glu_p	[x2**3**4x]
Protein	Arg β (d)	0.28	0.02	0.30	0.08	0.28	0.08	Glu_p	[x**23**4x] + [x2**34**x]
Protein	Arg β (t)	0.00	0.02	0.12	0.08	0.00	0.08	Glu_p	[x**234**x]
Protein	Arg δ (s)	0.56	0.01	0.42	0.06	0.59	0.04	Glu_p	[xxx4**5**]
Protein	Arg δ (d)	0.44	0.01	0.58	0.06	0.41	0.04	Glu_p	[xxx**45**]
Protein	Asp α (s)	0.30	0.04	0.37	0.04	0.28	0.04	OAA	[1**2**3x]
Protein	Asp α (d1)	0.12	0.04	0.17	0.04	0.16	0.04	OAA	[**12**3x]
Protein	Asp α (d2)	0.24	0.04	0.18	0.04	0.26	0.04	OAA	[1**23**x]
Protein	Asp α (dd)	0.34	0.04	0.28	0.04	0.29	0.04	OAA	[**123**x]
Protein	Asp β (s)	0.32	0.03	0.30	0.02	0.33	0.08	OAA	[x2**3**4]
Protein	Asp β (d1)	0.32	0.03	0.31	0.02	0.31	0.08	OAA	[x**23**4]
Protein	Asp β (d2)	0.18	0.03	0.25	0.02	0.20	0.08	OAA	[x2**34**]

Table S4: (continued.)

Protein	Asp β (dd)	0.19	0.03	0.15	0.02	0.16	0.08	OAA	[x**23**4]
Protein	Glu β (s)	0.58	0.02	0.69	0.04	0.70	0.04	Glu	[x2**3**4x]
Protein	Glu β (d)	0.42	0.02	0.28	0.04	0.27	0.04	Glu	[x**23**4x] + [x2**34**x]
Protein	Glu β (t)	0.00	0.02	0.03	0.04	0.03	0.04	Glu	[x**234**x]
Protein	Glu γ (s)	0.37	0.01	0.45	0.03	0.48	0.05	Glu	[xx3**4**5]
Protein	Glu γ (d1)	0.05	0.01	0.05	0.03	0.06	0.05	Glu	[xx**34**5]
Protein	Glu γ (d2)	0.56	0.01	0.46	0.03	0.44	0.05	Glu	[xx3**45**]
Protein	Glu γ (dd)	0.03	0.01	0.04	0.03	0.02	0.05	Glu	[xx**345**]
Protein	Gly α (s)	0.65	0.03	0.41	0.03	0.59	0.04	T3P/GOx	[1**2**x]
Protein	Gly α (d)	0.35	0.03	0.59	0.03	0.41	0.04	T3P/GOx	[**12**x]
Protein	His β (s)	0.57	0.06	0.35	0.05	0.30	0.05	$P5P_p$	[x2**3**4x]
Protein	His β (d1)	0.43	0.06	0.05	0.05	0.07	0.05	$P5P_p$	[x**23**4x]
Protein	His β (d2)	0.31	0.06	0.55	0.05	0.52	0.05	$P5P_p$	[x2**34**x]
Protein	His β (dd)	0.03	0.06	0.06	0.05	0.12	0.05	$P5P_p$	[x**234**x]
Protein	His δ2 (s)	0.55	0.09	0.28	0.10	0.57	0.08	$P5P_p$	[**1**2xxx]
Protein	His δ2 (d)	0.12	0.09	0.72	0.10	0.43	0.08	$P5P_p$	[**12**xxx]
Protein	Ile α (s)	0.52	0.01	0.45	0.06	0.48	0.06	OAA_p/Pyr_p	*[12xx]*·[x2x]
Protein	Ile α (d1)	0.31	0.01	0.30	0.06	0.34	0.06	OAA_p/Pyr_p	*[**12**xx]*·[x2x]
Protein	Ile α (d2)	0.04	0.01	0.09	0.06	0.04	0.06	OAA_p/Pyr_p	*[12xx]*·[x**2**x]
Protein	Ile α (dd)	0.13	0.01	0.16	0.06	0.15	0.06	OAA_p/Pyr_p	*[**12**xx]*·[x**2**x]

Hydroly-	**Cross peak**	**PGM-KO 24 culture 1**		**PGM-KO 24 culture 2**		**PGM-KO 24 culture 3**			
sate	**(multiplet)**	**RI**	**SD**	**RI**	**SD**	**RI**	**SD**	**Precursor**	**Isotopomer**
Protein	Ile γ1(s)	0.68	0.02	0.44	0.07	0.68	0.07	Pyr_p/OAA_p	[x2x]·*[xx**3**4]*
Protein	Ile γ1(d)	0.32	0.02	0.45	0.07	0.32	0.07	Pyr_p/OAA_p	[x**2**x]·*[xx**34**]* + [x2x]·*[xx**34**]*

Table S4: (continued.)

Protein	Ile γ1(t)	0.00	0.02	0.11	0.07	0.00	0.07	Pyr_p/OAA_p	[x**2**x]·*[xx**34**]*
Protein	Ile γ2 (s)	0.31	0.01	0.58	0.02	0.53	0.03	Pyr_p	[x2**3**]
Protein	Ile γ2 (d)	0.69	0.01	0.42	0.02	0.47	0.03	Pyr_p	[x**23**]
Protein	Ile δ (s)	0.66	0.01	0.55	0.07	0.34	0.08	OAA_p	[xx3**4**]
Protein	Ile δ (d)	0.34	0.01	0.45	0.07	0.66	0.08	OAA_p	[xx**34**]
Protein	Leu α (s)	0.24	0.07	0.26	0.04	0.27	0.05	$ACoA_p/Pyr_p$	*[**12**]*·[x2x]
Protein	Leu α (d1)	0.57	0.07	0.60	0.04	0.63	0.05	$ACoA_p/Pyr_p$	*[**12**]*·[x2x]
Protein	Leu α (d2)	0.12	0.07	0.04	0.04	0.05	0.05	$ACoA_p/Pyr_p$	*[**12**]*·[x**2**x]
Protein	Leu α (dd)	0.07	0.07	0.09	0.04	0.06	0.05	$ACoA_p/Pyr_p$	*[**12**]*·[x**2**x]
Protein	Leu b (s)	0.84	0.01	0.78	0.05	0.73	0.07	$ACoA_p/Pyr_p$	*[x2]*·[x**2**x].[x2x]
Protein	Leu b (d)	0.16	0.01	0.15	0.05	0.17	0.07	$ACoA_p/Pyr_p$	*[x**2**]*·[x**2**x].[x2x]+*[x2]·[x2x].[x2x]*
Protein	Leu b (t)	0.00	0.01	0.07	0.05	0.10	0.07	$ACoA_p/Pyr_p$	*[x**2**]*·[x**2**x].[x**2**x]
Protein	Leu δ1 (s)	0.45	0.07	0.43	0.03	0.31	0.03	Pyr_p	[x2**3**]
Protein	Leu δ1 (d)	0.55	0.07	0.57	0.03	0.69	0.03	Pyr_p	[x**23**]
Protein	Leu δ2 (s)	0.91	0.01	0.86	0.02	0.91	0.04	Pyr_p	[x2x]·[xx**3**]
Protein	Leu δ2 (d)	0.09	0.01	0.14	0.02	0.09	0.04	Pyr_p	[x**2**x]·[xx**3**]
Protein	Lys β (s)	0.26	0.01	0.22	0.08	0.34	0.05	OAA_p/Pyr_p	½{*[x2**34**]* + [x2**3**]·*[xxx4]*}
Protein	Lys β (d)	0.74	0.01	0.52	0.08	0.66	0.05	OAA_p/Pyr_p	½{*[x**234**]* + *[x2**34**]* +
								OAA_p/Pyr_p	[x**23**]·*[xxx4]* + [x2**3**]·*[xxx**4**]*}
Protein	Lys β (t)	0.00	0.01	0.26	0.08	0.00	0.05	OAA_p/Pyr_p	½{*[x**234**]* + [x**23**]·*[xxx**4**]*}
Protein	Lys γ (s)	0.59	0.03	0.57	0.09	0.47	0.04	OAA_p/Pyr_p	*[xx3**4**]*·[xx3]
Protein	Lys γ (d)	0.41	0.03	0.37	0.09	0.48	0.04	OAA_p/Pyr_p	*[xx**34**]*·[xx3] + *[xx3**4**]·[xx**3**]*
Protein	Lys γ (t)	0.00	0.03	0.06	0.09	0.05	0.04	OAA_p/Pyr_p	*[xx**34**]*·[xx**3**]
Protein	Lys δ (s)	0.31	0.01	0.32	0.07	0.25	0.08	OAA_p/Pyr_p	½{*[x2**34**]* + [x2**3**]·*[xxx4]*}

Table S4: (continued.)

Protein	Lys δ (d)	0.59	0.01	0.50	0.07	0.59	0.08	OAA_p/Pyr_p	½{*[x**234**]* + *[x2**34**]* + [x**23**]·*[xxx4]* + [x2**3**]·*[xxx**4**]*}
Protein	Lys δ (t)	0.10	0.01	0.18	0.07	0.16	0.08	OAA_p/Pyr_p	½{*[x**234**]* + [x**23**]·*[xxx**4**]*}
Protein	Lys ε (s)	0.41	0.01	0.47	0.05	0.38	0.06	OAA_p/Pyr_p	½{[x**2**3] + *[x**2**3x]*}
Protein	Lys ε (d)	0.59	0.01	0.53	0.05	0.62	0.06	OAA_p/Pyr_p	½{[x**23**] + *[x**23**x]*}
Protein	Phe α (s)	0.23	0.02	0.18	0.03	0.21	0.05	PEP_p	[1**2**3]
Protein	Phe α (d1)	0.10	0.02	0.02	0.03	0.07	0.05	PEP_p	[**12**3]
Protein	Phe α (d2)	0.10	0.02	0.35	0.03	0.12	0.05	PEP_p	[1**23**]
Protein	Phe α (dd)	0.57	0.02	0.46	0.03	0.60	0.05	PEP_p	[**123**]
Protein	Phe β (s)	0.24	0.02	0.17	0.07	0.29	0.03	PEP_p	[x2**3**]·[x2x]
Protein	Phe β (d1)	0.66	0.02	0.72	0.07	0.60	0.03	PEP_p	[x**23**]·[x2x]
Protein	Phe β (d2)	0.05	0.02	0.02	0.07	0.04	0.03	PEP_p	[x2**3**]·[x**2**x]
Protein	Phe β (dd)	0.04	0.02	0.09	0.07	0.07	0.03	PEP_p	[x**23**]·[x**2**x]
Protein	Pro γ (s)	0.50	0.13	0.29	0.07	0.41	0.04	Glu_p	[xx3**4**5]
Protein	Pro γ (d)	0.50	0.13	0.60	0.07	0.47	0.04	Glu_p	[xx**34**5] + [xx3**45**]
Protein	Pro γ (t)	0.00	0.13	0.12	0.07	0.12	0.04	Glu_p	[xx**345**]
Protein	Pro δ (s)	0.46	0.11	0.42	0.05	0.45	0.05	Glu_p	[xxx4**5**]
Protein	Pro δ (d)	0.54	0.11	0.58	0.05	0.55	0.05	Glu_p	[xxx**45**]

Hydroly-	**Cross peak**	**PGM-KO 24 culture 1**		**PGM-KO 24 culture 2**		**PGM-KO 24 culture 3**			
sate	**(multiplet)**	**RI**	**SD**	**RI**	**SD**	**RI**	**SD**	**Precursor**	**Isotopomer**
Protein	Ser α (s)	0.28	0.01	0.26	0.03	0.25	0.09	Ser	[1**2**3]
Protein	Ser α (d1)	0.18	0.01	0.19	0.03	0.19	0.09	Ser	[**12**3]
Protein	Ser α (d2)	0.09	0.01	0.14	0.03	0.13	0.09	Ser	[1**23**]

Table S4: (continued.)

Protein	Ser α (dd)	0.45	0.01	0.42	0.03	0.42	0.09	Ser	[**123**]
Protein	Ser β (s)	0.48	0.01	0.40	0.05	0.52	0.03	Ser	[x23**]**
Protein	Ser β (d)	0.52	0.01	0.60	0.05	0.48	0.03	Ser	[x**23**]
Protein	Thr α (s)	0.31	0.02	0.38	0.05	0.33	0.10	OAA_p	[1**2**3x]
Protein	Thr α (d1)	0.15	0.02	0.11	0.05	0.17	0.10	OAA_p	[**12**3x]
Protein	Thr α (d2)	0.27	0.02	0.27	0.05	0.25	0.10	OAA_p	[1**23**x]
Protein	Thr α (dd)	0.27	0.02	0.24	0.05	0.25	0.10	OAA_p	[**123**x]
Protein	Thr γ2 (s)	0.66	0.01	0.68	0.02	0.66	0.03	OAA_p	[xx3**4**]
Protein	Thr γ2 (d)	0.34	0.01	0.32	0.02	0.34	0.03	OAA_p	[xx**34**]
Protein	Tyr α (s)	0.23	0.01	0.24	0.07	0.25	0.08	PEP_p	[1**2**3]
Protein	Tyr α (d1)	0.05	0.01	0.10	0.07	0.08	0.08	PEP_p	[**12**3]
Protein	Tyr α (d2)	0.16	0.01	0.14	0.07	0.08	0.08	PEP_p	[1**23**]
Protein	Tyr α (dd)	0.57	0.01	0.52	0.07	0.59	0.08	PEP_p	[**123**]
Protein	Tyr β (s)	0.29	0.07	0.17	0.05	0.29	0.03	PEP_p	[x2**3**]·[x2x]
Protein	Tyr β (d)	0.71	0.07	0.74	0.05	0.64	0.03	PEP_p	[x**23**]·[x2x] + [x23]·[x2x]
Protein	Tyr β (t)	0.00	0.07	0.09	0.05	0.07	0.03	PEP_p	[x**23**]·[x**2**x]
Protein	Tyr δ1 (s)	0.34	0.03	0.16	0.06	0.29	0.03	PEP_p/$E4P_p$	*[x23]*·[1xxx]
Protein	Tyr δ1 (d)	0.66	0.03	0.75	0.06	0.71	0.03	PEP_p/$E4P_p$	*[x**23**]*·[1xxx] + *[x23]*·[**1**xxx]
Protein	Tyr δ1 (t)	0.00	0.03	0.09	0.06	0.00	0.03	PEP_p/$E4P_p$	*[x**23**]*·[**1**xxx]
Protein	Tyr ε1 (s)	0.59	0.05	0.40	0.04	0.54	0.04	PEP_p/$E4P_p$	*[xx3]*·[**1**2xx]
Protein	Tyr ε1 (d)	0.15	0.05	0.31	0.04	0.12	0.04	PEP_p/$E4P_p$	*[xx**3**]*·[**1**2xx] + *[xx3]*·[**12**xx]
Protein	Tyr ε1 (t)	0.27	0.05	0.29	0.04	0.35	0.04	PEP_p/$E4P_p$	*[xx**3**]*·[**12**xx]
Protein	Val α (s)	0.33	0.01	0.32	0.09	0.33	0.09	Pyr_p	[1**2**x]·[x2x]
Protein	Val α (d1)	0.44	0.01	0.43	0.09	0.42	0.09	Pyr_p	[**12**x]·[x2x]

Table S4: (continued.)

Protein	Val α (d2)	0.02	0.01	0.03	0.09	0.02	0.09	Pyr_p	[1**2**x]·[x**2**x]
Protein	Val α (dd)	0.22	0.01	0.22	0.09	0.23	0.09	Pyr_p	[1**2**x]·[x**2**x]
Protein	Val γ1 (s)	0.31	0.04	0.25	0.05	0.28	0.02	Pyr_p	[x2**3**]
Protein	Val γ1 (d)	0.69	0.04	0.75	0.05	0.72	0.02	Pyr_p	[x**23**]
Protein	Val γ2 (s)	0.90	0.01	0.82	0.06	0.91	0.02	Pyr_p	[x2x]·[xx**3**]
Protein	Val γ2 (d)	0.10	0.01	0.18	0.06	0.09	0.02	Pyr_p	[x**2**x]·[xx**3**]

Hydroly-	**Cross peak**	**PGM-KO 24 culture 1**		**PGM-KO 24 culture 2**		**PGM-KO 24 culture 3**			
sate	**(multiplet)**	**RI**	**SD**	**RI**	**SD**	**RI**	**SD**	**Precursor**	**Isotopomer**
Protein	HMF 1 (s)	0.24	0.11	0.25	0.15	0.18	0.08	$G6P_c/F6P_c$	[**1**2xxxx]
Protein	HMF 1 (d)	0.76	0.11	0.75	0.15	0.82	0.08	$G6P_c/F6P_c$	[**12**xxxx]
Protein	LVA 3 (s)	0.60	0.03	0.65	0.01	0.70	0.02	$G6P_c/F6P_c$	[x2**3**4xx]
Protein	LVA 3 (d1)	0.03	0.03	0.02	0.01	0.03	0.02	$G6P_c/F6P_c$	[x**23**4xx]
Protein	LVA 3 (d2)	0.01	0.03	0.00	0.01	0.00	0.02	$G6P_c/F6P_c$	[x2**34**xx]
Protein	LVA 3 (dd)	0.36	0.03	0.33	0.01	0.26	0.02	$G6P_c/F6P_c$	[x**234**xx]
Protein	LVA 6 (s)	0.11	0.01	0.15	0.01	0.17	0.01	$G6P_c/F6P_c$	[xxx45**6**]
Protein	LVA 6 (d1)	0.05	0.01	0.01	0.01	0.06	0.01	$G6P_c/F6P_c$	[xxx4**56**]
Protein	LVA 6 (d2)	0.21	0.01	0.25	0.01	0.13	0.01	$G6P_c/F6P_c$	[xxx**456**]
Protein	LVA 6 (dd)	0.63	0.01	0.58	0.01	0.64	0.01	$G6P_c/F6P_c$	[xxx**456**]

Table S5: Relative multiplet intensities of amino acids from protein and glucosyl units from starch hydrolysates with their standard deviations (SDs) from 2-D [^{13}C, ^{1}H] HSQC spectrum of PGM-KO 38 cultures. s indicates singlet, d1 and d2 indicate the first and second doublet and dd indicates the double doublet. Bold-faced carbon atom in Isotopomer indicates labeled carbon atom, normal font indicates unlabeled carbon atom and x indicates unknown labeling state of the carbon atom. Subscripts p, c and m for precursor indicates the compartments plastid, cytosol and mitochondrion in which they are present. RI and SD represent relative intensities and standard deviations respectively.

Hydroly-	Cross peak	PGM-KO 38 culture 1		PGM-KO 38 culture 2		PGM-KO 38 culture 3			
sate	(multiplet)	RI	SD	RI	SD	RI	SD	Precursor	Isotopomer
Protein	Ala α (s)	0.22	0.05	0.21	0.04	0.21	0.03	Pyr	[1**2**3]
Protein	Ala α (d1)	0.09	0.05	0.08	0.04	0.10	0.03	Pyr	[**12**3]
Protein	Ala α (d2)	0.52	0.05	0.59	0.04	0.59	0.03	Pyr	[1**23**]
Protein	Ala α (dd)	0.17	0.05	0.11	0.04	0.10	0.03	Pyr	[**123**]
Protein	Ala β (s)	0.08	0.06	0.26	0.02	0.24	0.02	Pyr	[x2**3**]
Protein	Ala β (d)	0.92	0.06	0.74	0.02	0.76	0.02	Pyr	[x**23**]
Protein	Arg β (s)	0.76	0.09	0.67	0.03	0.66	0.04	Glu_p	[x2**3**4x]
Protein	Arg β (d)	0.24	0.09	0.25	0.03	0.24	0.04	Glu_p	[x**23**4x] + [x2**34**x]
Protein	Arg β (t)	0.00	0.09	0.08	0.03	0.10	0.04	Glu_p	[x**234**x]
Protein	Arg δ (s)	0.53	0.04	0.33	0.04	0.52	0.02	Glu_p	[xxx4**5**]
Protein	Arg δ (d)	0.47	0.04	0.67	0.04	0.48	0.02	Glu_p	[xxx**45**]
Protein	Asp α (s)	0.35	0.02	0.36	0.05	0.33	0.03	OAA	[1**2**3x]
Protein	Asp α (d1)	0.16	0.02	0.20	0.05	0.19	0.03	OAA	[**12**3x]
Protein	Asp α (d2)	0.21	0.02	0.22	0.05	0.28	0.03	OAA	[1**23**x]
Protein	Asp α (dd)	0.18	0.02	0.23	0.05	0.20	0.03	OAA	[**123**x]
Protein	Asp β (s)	0.26	0.02	0.22	0.03	0.20	0.02	OAA	[x2**3**4]
Protein	Asp β (d1)	0.35	0.02	0.33	0.03	0.38	0.02	OAA	[x**23**4]
Protein	Asp β (d2)	0.33	0.02	0.38	0.03	0.35	0.02	OAA	[x2**34**]

Table S5: (continued.)

Protein	Asp β (dd)	0.06	0.02	0.07	0.03	0.07	0.02	OAA	[x**234**]
Protein	Glu β (s)	0.72	0.03	0.53	0.05	0.68	0.02	Glu	[x2**3**4x]
Protein	Glu β (d)	0.28	0.03	0.35	0.05	0.29	0.02	Glu	[x**23**4x] + [x2**34**x]
Protein	Glu β (t)	0.00	0.03	0.12	0.05	0.03	0.02	Glu	[x**234**x]
Protein	Glu γ (s)	0.48	0.02	0.47	0.04	0.44	0.02	Glu	[xx3**4**5]
Protein	Glu γ (d1)	0.07	0.02	0.01	0.04	0.05	0.02	Glu	[xx**34**5]
Protein	Glu γ (d2)	0.43	0.02	0.49	0.04	0.48	0.02	Glu	[xx3**45**]
Protein	Glu γ (dd)	0.02	0.02	0.02	0.04	0.03	0.02	Glu	[xx**345**]
Protein	Gly α (s)	0.51	0.04	0.49	0.04	0.53	0.02	T3P/GOx	[1**2**x]
Protein	Gly α (d)	0.49	0.04	0.51	0.04	0.47	0.02	T3P/GOx	[**12**x]
Protein	His β (s)	0.34	0.04	0.35	0.06	0.33	0.07	$P5P_p$	[x2**3**4x]
Protein	His β (d1)	0.09	0.04	0.05	0.06	0.03	0.07	$P5P_p$	[x**23**4x]
Protein	His β (d2)	0.51	0.04	0.55	0.06	0.54	0.07	$P5P_p$	[x2**34**x]
Protein	His β (dd)	0.06	0.04	0.06	0.06	0.10	0.07	$P5P_p$	[x**234**x]
Protein	His δ2 (s)	0.58	0.05	0.55	0.06	0.57	0.07	$P5P_p$	[**1**2xxx]
Protein	His δ2 (d)	0.42	0.05	0.45	0.06	0.43	0.07	$P5P_p$	[**12**xxx]
Protein	Ile α (s)	0.47	0.07	0.51	0.07	0.50	0.04	OAA_p/Pyr_p	*[**12**xx]*·[x2x]
Protein	Ile α (d1)	0.28	0.07	0.32	0.07	0.26	0.04	OAA_p/Pyr_p	*[**12**xx]*·[x2x]
Protein	Ile α (d2)	0.09	0.07	0.04	0.07	0.04	0.04	OAA_p/Pyr_p	*[12xx]*·[x**2**x]
Protein	Ile α (dd)	0.16	0.07	0.14	0.07	0.19	0.04	OAA_p/Pyr_p	*[**12**xx]*·[x**2**x]

Hydroly-	**Cross peak**	**PGM-KO 38 culture 1**		**PGM-KO 38 culture 2**		**PGM-KO 38 culture 3**			
sate	**(multiplet)**	**RI**	**SD**	**RI**	**SD**	**RI**	**SD**	**Precursor**	**Isotopomer**
Protein	Ile γ1(s)	0.70	0.06	0.51	0.05	0.39	0.05	Pyr_p/OAA_p	[x2x]·*[xx**34**]*

Table S5: (continued.)

Protein	Ile γ1(d)	0.30	0.06	0.45	0.05	0.48	0.05	Pyr_p/OAA_p	[x**2**x]·*[xx**3**4]* + [x2x]·*[xx**34**]*
Protein	Ile γ1(t)	0.00	0.06	0.05	0.05	0.13	0.05	Pyr_p/OAA_p	[x**2**x]·*[xx**34**]*
Protein	Ile γ2 (s)	0.69	0.04	0.53	0.03	0.70	0.02	Pyr_p	[x2**3**]
Protein	Ile γ2 (d)	0.31	0.04	0.47	0.03	0.30	0.02	Pyr_p	[x**23**]
Protein	Ile δ (s)	0.40	0.04	0.38	0.06	0.34	0.06	OAA_p	[xx3**4**]
Protein	Ile δ (d)	0.60	0.04	0.62	0.06	0.66	0.06	OAA_p	[xx**34**]
Protein	Leu α (s)	0.07	0.06	0.28	0.04	0.28	0.05	$ACoA_p/Pyr_p$	*[**12**]*·[x2x]
Protein	Leu α (d1)	0.04	0.06	0.62	0.04	0.62	0.05	$ACoA_p/Pyr_p$	*[**12**]*·[x2x]
Protein	Leu α (d2)	0.18	0.06	0.03	0.04	0.04	0.05	$ACoA_p/Pyr_p$	*[**12**]*·[x**2**x]
Protein	Leu α (dd)	0.71	0.06	0.07	0.04	0.06	0.05	$ACoA_p/Pyr_p$	*[**12**]*·[x**2**x]
Protein	Leu b (s)	0.90	0.03	0.80	0.03	0.90	0.03	$ACoA_p/Pyr_p$	*[x2]*·[x**2**x].[x2x]
Protein	Leu b (d)	0.10	0.03	0.19	0.03	0.04	0.03	$ACoA_p/Pyr_p$	*[x**2**]*·[x**2**x].[x2x]+*[x2]·[x2x].[x2x]*
Protein	Leu b (t)	0.00	0.03	0.01	0.03	0.06	0.03	$ACoA_p/Pyr_p$	*[x**2**]*·[x**2**x].[x**2**x]
Protein	Leu δ1 (s)	0.44	0.01	0.43	0.02	0.41	0.02	Pyr_p	[x2**3**]
Protein	Leu δ1 (d)	0.56	0.01	0.57	0.02	0.59	0.02	Pyr_p	[x**23**]
Protein	Leu δ2 (s)	0.92	0.01	0.91	0.02	0.91	0.01	Pyr_p	[x2x]·[xx**3**]
Protein	Leu δ2 (d)	0.08	0.01	0.09	0.02	0.09	0.01	Pyr_p	[x**2**x]·[xx**3**]
Protein	Lys β (s)	0.23	0.03	0.33	0.04	0.26	0.02	OAA_p/Pyr_p	½{*[x2**3**4]* + [x2**3**]·*[xxx4]*}
Protein	Lys β (d)	0.65	0.03	0.55	0.04	0.65	0.02	OAA_p/Pyr_p	½{*[x**23**4]* + *[x2**34**]* +
								OAA_p/Pyr_p	[x**23**]·*[xxx4]* + [x2**3**]·*[xxx**4**]*}
Protein	Lys β (t)	0.12	0.03	0.12	0.04	0.10	0.02	OAA_p/Pyr_p	½{*[x**234**]* + [x**23**]·*[xxx**4**]*}
Protein	Lys γ (s)	0.62	0.05	0.52	0.06	0.46	0.04	OAA_p/Pyr_p	*[xx3**4**]*·[xx3]
Protein	Lys γ (d)	0.38	0.05	0.35	0.06	0.33	0.04	OAA_p/Pyr_p	*[xx**34**]*·[xx3] + *[xx3**4**]·[xx**3**]*
Protein	Lys γ (t)	0.00	0.05	0.13	0.06	0.21	0.04	OAA_p/Pyr_p	*[xx**34**]*·[xx**3**]

Table S5: (continued.)

Protein	Lys δ (s)	0.27	0.05	0.19	0.05	0.34	0.03	OAA_p/Pyr_p	½{*[x2**34**]* + [x2**3**]·*[xxx4]*}
Protein	Lys δ (d)	0.58	0.05	0.63	0.05	0.52	0.03	OAA_p/Pyr_p	½{*[x**234**]* + *[x2**34**]* + [x**23**]·*[xxx4]* + [x2**3**]·*[xxx**4**]*}
Protein	Lys δ (t)	0.16	0.05	0.18	0.05	0.13	0.03	OAA_p/Pyr_p	½{*[x**234**]* + [x**23**]·*[xxx**4**]*}
Protein	Lys ε (s)	0.31	0.02	0.43	0.05	0.39	0.01	OAA_p/Pyr_p	½{[x**2**3] + *[x**2**3x]*}
Protein	Lys ε (d)	0.69	0.02	0.57	0.05	0.61	0.01	OAA_p/Pyr_p	½{[x**23**] + *[x**23**x]*}
Protein	Phe α (s)	0.24	0.07	0.24	0.04	0.24	0.03	PEP_p	[1**2**3]
Protein	Phe α (d1)	0.07	0.07	0.09	0.04	0.05	0.03	PEP_p	[**12**3]
Protein	Phe α (d2)	0.15	0.07	0.13	0.04	0.14	0.03	PEP_p	[1**23**]
Protein	Phe α (dd)	0.54	0.07	0.54	0.04	0.57	0.03	PEP_p	[**123**]
Protein	Phe β (s)	0.33	0.04	0.31	0.06	0.34	0.02	PEP_p	[x2**3**]·[x2x]
Protein	Phe β (d1)	0.34	0.04	0.30	0.06	0.55	0.02	PEP_p	[x**23**]·[x2x]
Protein	Phe β (d2)	0.20	0.04	0.22	0.06	0.04	0.02	PEP_p	[x2**3**]·[x**2**x]
Protein	Phe β (dd)	0.13	0.04	0.17	0.06	0.07	0.02	PEP_p	[x**23**]·[x**2**x]
Protein	Pro γ (s)	0.22	0.07	0.43	0.05	0.39	0.02	Glu_p	[xx3**4**5]
Protein	Pro γ (d)	0.64	0.07	0.47	0.05	0.49	0.02	Glu_p	[xx**34**5] + [xx3**45**]
Protein	Pro γ (t)	0.14	0.07	0.10	0.05	0.12	0.02	Glu_p	[xx**345**]
Protein	Pro δ (s)	0.45	0.07	0.48	0.02	0.47	0.04	Glu_p	[xxx4**5**]
Protein	Pro δ (d)	0.55	0.07	0.52	0.02	0.53	0.04	Glu_p	[xxx**45**]

Hydroly-	**Cross peak**	**PGM-KO 38 culture 1**		**PGM-KO 38 culture 2**		**PGM-KO 38 culture 3**			
sate	**(multiplet)**	**RI**	**SD**	**RI**	**SD**	**RI**	**SD**	**Precursor**	**Isotopomer**
Protein	Ser α (s)	0.28	0.05	0.24	0.04	0.25	0.03	Ser	[1**2**3]
Protein	Ser α (d1)	0.18	0.05	0.19	0.04	0.16	0.03	Ser	[**12**3]

Table S5: (continued.)

Protein	Ser α (d2)	0.13	0.05	0.11	0.04	0.14	0.03	Ser	[1**23**]
Protein	Ser α (dd)	0.41	0.05	0.46	0.04	0.45	0.03	Ser	[**123**]
Protein	Ser β (s)	0.55	0.05	0.54	0.04	0.53	0.02	Ser	[x2**3**]
Protein	Ser β (d)	0.45	0.05	0.46	0.04	0.47	0.02	Ser	[x**23**]
Protein	Thr α (s)	0.36	0.04	0.31	0.07	0.31	0.05	OAA_p	[12**3**x]
Protein	Thr α (d1)	0.15	0.04	0.12	0.07	0.12	0.05	OAA_p	[**12**3x]
Protein	Thr α (d2)	0.22	0.04	0.31	0.07	0.28	0.05	OAA_p	[1**23**x]
Protein	Thr α (dd)	0.26	0.04	0.26	0.07	0.29	0.05	OAA_p	[**123**x]
Protein	Thr γ2 (s)	0.71	0.01	0.65	0.02	0.69	0.03	OAA_p	[xx3**4**]
Protein	Thr γ2 (d)	0.29	0.01	0.35	0.02	0.31	0.03	OAA_p	[xx**34**]
Protein	Tyr α (s)	0.21	0.08	0.21	0.04	0.21	0.03	PEP_p	[1**2**3]
Protein	Tyr α (d1)	0.08	0.08	0.08	0.04	0.09	0.03	PEP_p	[**12**3]
Protein	Tyr α (d2)	0.10	0.08	0.15	0.04	0.14	0.03	PEP_p	[1**23**]
Protein	Tyr α (dd)	0.62	0.08	0.56	0.04	0.57	0.03	PEP_p	[**123**]
Protein	Tyr β (s)	0.29	0.04	0.31	0.06	0.20	0.03	PEP_p	[x2**3**]·[x2x]
Protein	Tyr β (d)	0.71	0.04	0.69	0.06	0.55	0.03	PEP_p	[x**23**]·[x2x] + [x23]·[x2x]
Protein	Tyr β (t)	0.00	0.04	0.00	0.06	0.25	0.03	PEP_p	[x**23**]·[x**2**x]
Protein	Tyr δ1 (s)	0.48	0.02	0.30	0.03	0.10	0.06	PEP_p/$E4P_p$	*[x2**3**]*·[1xxx]
Protein	Tyr δ1 (d)	0.52	0.02	0.71	0.03	0.90	0.06	PEP_p/$E4P_p$	*[x**23**]*·[1xxx] + *[x2**3**]*·[**1**xxx]
Protein	Tyr δ1 (t)	0.00	0.02	0.00	0.03	0.00	0.06	PEP_p/$E4P_p$	*[x**23**]*·[**1**xxx]
Protein	Tyr ε1 (s)	0.56	0.05	0.60	0.03	0.57	0.03	PEP_p/$E4P_p$	*[xx3]*·[**1**2xx]
Protein	Tyr ε1 (d)	0.18	0.05	0.13	0.03	0.12	0.03	PEP_p/$E4P_p$	*[xx**3**]*·[**1**2xx] + *[xx3]*·[**12**xx]
Protein	Tyr ε1 (t)	0.26	0.05	0.27	0.03	0.31	0.03	PEP_p/$E4P_p$	*[xx**3**]*·[**12**xx]
Protein	Val α (s)	0.34	0.05	0.33	0.03	0.37	0.02	Pyr_p	[1**2**x]·[x2x]

Table S5: (continued.)

Protein	Val α (d1)	0.41	0.05	0.39	0.03	0.52	0.02	Pyr_p	[**12**x]·[x2x]
Protein	Val α (d2)	0.02	0.05	0.02	0.03	0.04	0.02	Pyr_p	[1**2**x]·[x**2**x]
Protein	Val α (dd)	0.23	0.05	0.26	0.03	0.06	0.02	Pyr_p	[1**2**x]·[x**2**x]
Protein	Val γ1 (s)	0.37	0.01	0.32	0.02	0.43	0.03	Pyr_p	[x2**3**]
Protein	Val γ1 (d)	0.63	0.01	0.68	0.02	0.57	0.03	Pyr_p	[x**23**]
Protein	Val γ2 (s)	0.89	0.01	0.93	0.02	0.85	0.04	Pyr_p	[x2x]·[xx**3**]
Protein	Val γ2 (d)	0.11	0.01	0.07	0.02	0.15	0.04	Pyr_p	[x**2**x]·[xx**3**]

Hydroly-	**Cross peak**	**PGM-KO 38 culture 1**		**PGM-KO 38 culture 2**		**PGM-KO 38 culture 3**			
sate	**(multiplet)**	**RI**	**SD**	**RI**	**SD**	**RI**	**SD**	**Precursor**	**Isotopomer**
Protein	HMF 1 (s)	0.40	0.06	0.10	0.03	0.27	0.03	$G6P_c/F6P_c$	[**1**2xxxx]
Protein	HMF 1 (d)	0.60	0.06	0.90	0.03	0.73	0.03	$G6P_c/F6P_c$	[**12**xxxx]
Protein	LVA 3 (s)	0.72	0.02	0.67	0.02	0.83	0.01	$G6P_c/F6P_c$	[x2**3**4xx]
Protein	LVA 3 (d1)	0.18	0.02	0.22	0.02	0.11	0.01	$G6P_c/F6P_c$	[x**23**4xx]
Protein	LVA 3 (d2)	0.09	0.02	0.08	0.02	0.05	0.01	$G6P_c/F6P_c$	[x2**34**xx]
Protein	LVA 3 (dd)	0.01	0.02	0.03	0.02	0.01	0.01	$G6P_c/F6P_c$	[x**234**xx]
Protein	LVA 6 (s)	0.19	0.01	0.09	0.02	0.06	0.01	$G6P_c/F6P_c$	[xxx45**6**]
Protein	LVA 6 (d1)	0.02	0.01	0.01	0.02	0.00	0.01	$G6P_c/F6P_c$	[xxx4**56**]
Protein	LVA 6 (d2)	0.52	0.01	0.64	0.02	0.80	0.01	$G6P_c/F6P_c$	[xxx**4**5**6**]
Protein	LVA 6 (dd)	0.26	0.01	0.26	0.02	0.14	0.01	$G6P_c/F6P_c$	[xxx**456**]

CHAPTER 5

CONCLUSION

The results from *in vitro* culture demonstrated that the relative growth rate, sucrose and glutamine consumption rates, and biomass composition of SHaM embryos were dictated by the concentration of carbon and organic nitrogen sources in the media. A current understanding of the biochemistry of legumes, is that plants create a homeostatic state in the developing seed, by supplying a constant supply of carbon from photosynthesis in the source tissues. When the carbon supply to the developing zygote is in excess of that required to drive growth and maintenance metabolism, it is used for the biosynthesis of storage reserves. Overall, the results of these experiments revealed that protein biosynthesis was sensitive to decreasing C: N ratios in the culture media. Conversely, starch and residual biomass content increased with increasing C: N mole ratios. Embryo oil content however, appeared to be insensitive to the culture conditions. A better understanding of how carbon flux between protein and oil synthesis is regulated might lead to the development of molecular strategies to improve soybean seed quality. Metabolic flux analysis is an ideal experimental tool for studying such interactions, as it not only enables the flux into various pools to be estimated but it also brings insight into the contributions and interactions of the various metabolic compartments. The data presented here informs on the culture conditions that should provide the greatest insights into the metabolic control points of assimilate partitioning in SHaM embryos and by extension, developing soybean seeds.

A metabolic flux analysis with incorporation of transcript profiles of zygotic and SHaM embryos was performed using steady state ^{13}C MFA labeled sucrose. The reported flux values between the SHaM embryos and zygotic embryos differed, and these differences related to: sucrose and glutamine uptake fluxes, product synthesis (mainly oil), and the central metabolism of glycolysis, oxidative pentose phosphate pathway and the TCA cycle. However, the flux distributions in the three compartments of the model all follow the same trend. In addition, the biomass composition of starch and relative intensities of proteinogenic amino acids and glucosyl units in hydrolyzed starch in both embryos are very similar. Our results show that flux through the TCA cycle has an effect on the protein biosynthesis. This suggests that genetic manipulation through TCA cycle could lead to increase protein production in soybean. As the results show, SHaM embryos present a very suitable experimental model system for developing seed. Since soybean compositions are dictated by genetics and environment, one can use SHaM embryos as a fast and effective way to screen for genes target for future manipulation. Lastly, to narrow the target genes list that could be responsible for increase protein and oil production, we should integrate other global analytical methods such as transcriptomics, metabolimics, and proteomics.

A metabolic flux analysis using steady state ^{13}C MFA labeled sucrose can be applied to SHaM embryos in transgenic cultures, where the plastidic phosphoglucomutase (PGM) gene knocked out (PGM-KO) can be compared to a control (PGM-null). The results indicate that protein and oil contents of PGM-KO were consistently higher than the PGM-null. The synthesis of sucrose in the cytosolic

pathway was increased for PGM-KO, because there is no starch synthesis in the plastidic pathway. Notably, the fluxes from cytosolic transketolase, transaldolase, and glucose-6-phosphate isomerase increased in the absence of starch synthesis in the system. Furthermore, in PGM-KO culture, one of the malate transporters is used to regulate the flux from the TCA cycle into the cytosol to support the high oil and protein synthesis. Our results show that most of the fluxes are more affected by transgenic manipulation than by media alteration. This suggests that genetic manipulation is the major motivation for changing storage composition in the soybean embryos.